



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:55 ; Search time 6.67161 Seconds  
(without alignments)  
1409.756 Million cell updates/sec

Title: US-10-083-336A-10  
Perfect score: 1029  
Sequence: 1 MIFPKQYPIINFTTAGTAVQ.....RFQYIEGEMRTIRYNRRSA 200

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1024	99.5	576	1	RICI_RICCO
2	934.5	90.8	564	1	AGGL_RICCO
3	253.5	34.4	562	1	ABRC_ABRPR
4	345	33.5	527	1	ABRB_ABRPR
5	342	33.2	289	1	RIP2_TRIKI
6	341.5	33.2	282	1	RIP2_BRYDI
7	338	32.8	289	1	RIPS_TRIKI
8	330.5	32.1	563	1	NIGB_SAWNI
9	329.5	32.0	528	1	ABRA_ABRPR
10	312	30.3	290	1	RIP1_BRYDI
11	300.5	29.2	250	1	RIPB_LUFYCI
12	296.5	28.8	286	1	RIP1_CUCFI
13	286	27.8	254	1	MLA_VISAL
14	284.5	27.6	277	1	RIP4_LUFYCI
15	279	27.1	286	1	RIP2_MOMBA
16	277	26.9	286	1	RIP1_MOMCH
17	260	25.3	294	1	RIP1_TRIAN
18	245.5	23.9	316	1	RIPG_GELMU
19	197.5	19.2	294	1	RIP1_PHYAM
20	182	17.7	313	1	RIP1_PHYAM
21	180.5	17.5	261	1	RIPS_PHYAM
22	178	17.3	278	1	RIPP_MIRTA
23	140	13.6	280	1	RIP2_HORVU
24	133	12.9	280	1	RIP1_HORVU
25	131	12.7	253	1	RIP7_SAPOF
26	127	12.3	253	1	RIP5_SAPOF
27	126	12.2	310	1	RIP2_PHYAM
28	124	12.1	299	1	RIP6_SAPOF
29	123.5	12.0	293	1	RIP0_DIAAC
30	121	11.8	292	1	RIP2_SAPOF
31	111	10.8	319	1	SLTA_BP933
32	110	10.7	236	1	RIP3_SAPOF
33	109.5	10.6	300	1	RIP3_MAIZE

RESULT 1

RICI_RICCO	STANDARD;	PRT;	576 AA.
ID	RICI_RICCO		
AC	P02879; P02880;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase)		
DE	(BC 3.2.2.22); Ricin B chain].		
OS	Ricinus communis (Castor bean).		
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid I; Malpighiales; Euphorbiaceae; Ricinus.		
OX	NCBI_TaxID=3988;		
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	MEDLINE=86067214; PubMed=2999712;		
RX	Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,		
RA	Weaver R.F.;		
RA	"Genomic cloning and characterization of a ricin gene from Ricinus		
RT	communis";		
RL	Nucleic Acids Res. 13:8019-8033 (1985).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RP	MEDLINE=92163016; PubMed=1371405;		
RX	Tregear J.W., Roberts L.M.;		
RA	"The lectin gene family of Ricinus communis: cloning of a functional		
RT	ricin gene and three lectin pseudogenes.";		
RL	Plant Mol. Biol. 18:515-525 (1992).		
RL	[3]		
RN	SEQUENCE OF 12-576 FROM N.A.		
RP	MEDLINE=85179479; PubMed=3838723;		
RX	Lamb A., Roberts L.M., Lord J.M.;		
RA	"Nucleotide sequence of cloned cDNA coding for preproricin.";		
RT	Eur. J. Biochem. 148:265-270 (1985).		
RL	[4]		
RN	SEQUENCE OF 36-302.		
RP	Yoshitake S., Funatsu G., Funatsu M.;		
RA	"Isolation and sequences of peptic peptides, and the complete		
RT	sequence of ile chain of ricin-D.";		
RL	Agric. Biol. Chem. 42:1267-1274 (1978).		
RL	[5]		
RN	SEQUENCE OF 315-576.		
RP	Funatsu G., Kimura M., Funatsu M.;		
RA	"Primary structure of Ala chain of ricin D.";		
RT	Agric. Biol. Chem. 43:2221-2224 (1979).		
RL	[6]		
RN	CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.		
RP	MEDLINE=90344223; PubMed=1368517;		
RX	Kimura Y., Kusuoku H., Tada M., Takagi S., Funatsu G.;		
RA	"Structural analyses of sugar chains from ricin A-chain variant.";		
RT	Agric. Biol. Chem. 54:157-162 (1990).		
RL	[7]		
RN	REVIEW.		
RP	MEDLINE=21480122; PubMed=11595634;		
RX	Olsnes S., Kozlov J.V.;		
RA			

ALIGNMENTS

STANDARD; PRT; 576 AA.

DT 21-JUL-1986 (Rel. 01, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ricin precursor [Contains: Ricin A chain (xRNA N-glycosidase)  
(EC 3.2.2.22); Ricin B chain].  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RX MEDLINE=86067214; PubMed=2999712;  
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,  
RA Weaver R.F.;  
RT "Genomic cloning and characterization of a ricin gene from Ricinus  
communis";  
RL Nucleic Acids Res. 13:8019-8033(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92163016; PubMed=1371405;  
RA Tregear J.W., Roberts L.M.;  
RT "The lectin gene family of Ricinus communis: cloning of a functional  
ricin gene and three lectin pseudogenes";  
RL Plant Mol. Biol. 18:515-525(1992).  
RN [3]  
RP SEQUENCE OF 12-576 FROM N.A.  
RX MEDLINE=85179479; PubMed=3838723;  
RA Lamb A., Roberts L.M., Lord J.M.;  
RT "Nucleotide sequence of cloned cDNA coding for preproricin.";  
RL Eur. J. Biochem. 148:265-270(1985).  
RN [4]  
RP SEQUENCE OF 36-302.  
RA Yoshitake S., Funatsu G., Funatsu M.;  
RT "Isolation and sequences of peptic peptides, and the complete  
sequence of Ile chain of ricin-D";  
RL Agric. Biol. Chem. 42:1267-1274(1978).  
RN [5]  
RP SEQUENCE OF 315-576.  
RA Funatsu G., Kimura M., Funatsu M.;  
RT "Primary structure of Ala chain of ricin D";  
RL Agric. Biol. Chem. 43:2221-2224(1979).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.  
RX MEDLINE=90344223; PubMed=1368517;  
RA Kimura Y., Kusuoku H., Tada M., Takagi S., Funatsu G.;  
RT "Structural analyses of sugar chains from ricin A-chain variant.";  
RL Agric. Biol. Chem. 54:157-162(1990).  
RN [7]  
RP REVIEW.  
RX MEDLINE=21480122; PubMed=11595634;  
RA Olsnes S., Kozlov J.V.;

RT "Ricin.";  
 RL Toxicon 39:1723-1728(2001).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=87165983; PubMed=3558397;  
 RA Monfort W., Villafraña J.B., Monzingo A.F., Ernst S.R., Katzin B.,  
 RA Rutenber E., Kuong N.H., Hamlin R., Robertus J.D.;  
 RT "The three-dimensional structure of ricin at 2.8 Å.";  
 RL J. Biol. Chem. 262:5398-5403(1987).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=91352004; PubMed=1881881;  
 RA Katzin B.J., Collins E.J., Robertus J.D.;  
 RT "Structure of ricin A-chain at 2.5 Å.";  
 RL Proteins 10:251-259(1991).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
 RX MEDLINE=91352005; PubMed=1881882;  
 RA Rutenber E., Robertus J.D.;  
 RT "Structure of ricin B-chain at 2.5-Å resolution.";  
 RL Proteins 10:260-269(1991).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=95082010; PubMed=7590130;  
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,  
 RA Pauptit R.A.;  
 RT "X-ray structure of recombinant ricin A-chain at 1.8-Å resolution.";  
 RL J. Mol. Biol. 244:410-422(1994).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.  
 RX MEDLINE=96374222; PubMed=8780513;  
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,  
 RA Molina-Svinth M.C., Robertus J.D.;  
 RT "Structure and activity of an active site substitution of ricin A chain.";  
 RL Biochemistry 35:11098-11103(1996).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=97240820; PubMed=9086280;  
 RA Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Milne G.W.,  
 RA Robertus J.D.;  
 RT "Structure-based identification of a ricin inhibitor.";  
 RL J. Mol. Biol. 266:1043-1049(1997).  
 RN [14]  
 RP MUTAGENESIS.  
 RX MEDLINE=93165632; PubMed=1287657;  
 RA Kin Y., Robertus J.D.;  
 RT "Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography.";  
 RL Protein Eng. 5:775-779(1992).  
 CC -|- FUNCTION: Ricin is highly toxic to animal cells and to a less  
 CC extent to plant cells. The A chain is responsible for inhibiting  
 CC protein synthesis through the catalytic inactivation of 60S  
 CC ribosomal subunits. It acts as a glycosylase that removes a  
 CC specific adenine residue from an exposed loop of 28S ribosomal  
 CC RNA. As this loop is involved in the binding of elongation  
 CC factors, the modified ribosomes are unable to support protein  
 CC synthesis. The A chain can inactivate a few thousand ribosomes  
 CC per minute, thus inactivating them faster than the cell can make  
 CC new ones. A single A-chain molecule can therefore kill an animal  
 CC cell. The B chain binds to cell receptors and facilitates the  
 CC entry into the cell of the A chain; B chains are also responsible  
 CC for cell agglutination (lectin activity). It binds to beta-D-  
 CC galactopyranoside moieties.  
 CC -|- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -|- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 CC -|- DOMAIN: The B chain is composed of two domains, each domain  
 CC consists of 3 homologous subdomains (alpha, beta, gamma).  
 CC -|- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE  
 CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.  
 CC -|- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

CC -|- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -|- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS  
 CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).  
 CC -|- DATABASE: NAME=Protein Spotlight;  
 CC NOTE=Issue 31 of February 2003;  
 CC WWW="http://www.expasy.org/spotlight/articles/splt031.html".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X03179; CAA26939.1; -;  
 CC EMBL; X52908; CAA37095.1; -;  
 CC EMBL; X02388; CAA26230.1; -;  
 CC EMBL; A12892; CAA01058.1; -;  
 CC PIR; A24041; RLCSO.  
 CC PDB; 2AAI; 31-JAN-94.  
 CC PDB; 1APG; 31-JAN-94.  
 CC PDB; 1FMP; 31-OCT-93.  
 CC PDB; 1IFS; 14-JAN-98.  
 CC PDB; 1IFT; 14-JAN-98.  
 CC PDB; 1IFU; 14-JAN-98.  
 CC PDB; 1IKC; 31-OCT-93.  
 CC PDB; 1OBS; 16-JUN-97.  
 CC PDB; 1OBT; 16-JUN-97.  
 CC PDB; 1BR5; 02-SEP-98.  
 CC PDB; 1BR6; 02-SEP-98.  
 CC PDB; 1IL3; 16-JAN-02.  
 CC PDB; 1IL4; 16-JAN-02.  
 CC PDB; 1IL9; 16-JAN-02.  
 CC GlycoSuiteDB; P02879; -;  
 CC InterPro; IPR000772; Ricin\_B\_lectin.  
 CC InterPro; IPR001574; RIP.  
 CC Pfam; PF00652; Ricin\_B\_lectin; 6.  
 CC PRINTS; PR00396; SHIGARICIN.  
 CC SMART; SM00458; RICIN; 2.  
 CC PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
 CC PROSITE; PS00275; SHIGA\_RICIN; 1.  
 CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 CC Glycoprotein; Lectin; Signal; 3D-structure.  
 CC SIGNAL 1 35  
 CC CHAIN 36 302  
 CC PEPTIDE 303 314  
 CC CHAIN 315 576  
 CC DOMAIN 321 448  
 CC DOMAIN 451 575  
 CC REPEAT 331 373  
 CC REPEAT 374 414  
 CC REPEAT 417 449  
 CC REPEAT 462 497  
 CC REPEAT 501 540  
 CC REPEAT 543 570  
 CC ACT\_SITE 212 212  
 CC DISULFID 294 318  
 CC DISULFID 334 353  
 CC DISULFID 377 394  
 CC DISULFID 465 478  
 CC DISULFID 504 521  
 CC CARBOHYD 45 45  
 CC CARBOHYD 271 271  
 CC CARBOHYD 409 409  
 CC CARBOHYD 449 449  
 CC CARBOHYD 76 76  
 CC CONFLICT 551 551  
 CC STRAND 43 47  
 CC TURN 49 50  
 CC N-LINKED (GLCNAC. . .).  
 CC /FTIG-CAR 000080.  
 CC N-LINKED (GLCNAC. . .) (IN MINOR FORM).  
 CC /FTID-CAR 000081.  
 CC N-LINKED (GLCNAC. . .).  
 CC N-LINKED (GLCNAC. . .).  
 CC E -> D (IN REF. 3).  
 CC A -> R (IN REF. 3).  
 CC

Query Match 99.5%; Score 1024; DB 1; Length 576;  
Best Local Similarity 100.0%; Pred. No. 1.1e-87;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFPKQYPIINFTAGATVQSYNFIKAVRGRLLTGADVRHEIPVLPNVRVGLPINQRFILV 61  
DB 36 IFPKQYPIINFTAGATVQSYNFIKAVRGRLLTGADVRHEIPVLPNVRVGLPINQRFILV 95

QY 62 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDQDAEAIHTLFTDVQNRVYTFAG 121  
DB 96 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDQDAEAIHTLFTDVQNRVYTFAG 155

QY 122 GNYDRLEQLAGNLENIELGNGPLEBAISALYYSTGGTQPLTARSFFIOMISEAR 181  
DB 156 GNYDRLEQLAGNLENIELGNGPLEBAISALYYSTGGTQPLTARSFFIOMISEAR 215

QY 182 FOYIEGEMTRIRYNRRA 200  
DB 216 FOYIEGEMTRIRYNRRA 234

RESULT 2  
AGGL\_RICCO STANDARD; PRT; 564 AA.

AC P06750;  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
DE Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86059449; PubMed=2999130;  
RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;  
RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";  
RL J. Biol. Chem. 260:15682-15686(1985).  
RN [2]  
RP SEQUENCE OF 303-564.  
RC TISSUE=Seed;  
RA Araki T., Yoshioka Y., Funatsu G.;  
RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";  
RL Biochim. Biophys. Acta 872:277-285(1986).  
RN [3]  
RP SEQUENCE OF 303-337.  
RX MEDLINE=80178723; PubMed=6768555;  
RA Lin T.-S., Li S.-L.;  
RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";  
RL Eur. J. Biochem. 105:453-459(1980).  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenine on the 28S rRNA.  
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUPERFAMILY.  
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M12089; AAA33869.1; -;  
CC EMBL; S40368; AAB22584.1; -;

DR PIR; A24261; RLCSAG.  
DR HSP; P02879; LR86.  
DR GlycoSuiteDB; P06750; -;  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
KW Glycoprotein; Lectin; Signal.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 290 AGGLUTININ A CHAIN.  
FT PROPEP 291 302 LINKER PEPTIDE.  
FT CHAIN 303 564 AGGLUTININ B CHAIN.  
FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.  
FT REPEAT 319 361 1-ALPHA.  
FT REPEAT 362 402 1-BETA.  
FT REPEAT 405 437 1-GAMMA.  
FT REPEAT 450 485 2-ALPHA.  
FT REPEAT 489 528 2-BETA.  
FT REPEAT 531 558 2-GAMMA.  
FT ACT\_SITE 200 200 BY SIMILARITY.  
FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 322 341 BY SIMILARITY.  
FT DISULFID 365 382 BY SIMILARITY.  
FT DISULFID 453 466 BY SIMILARITY.  
FT DISULFID 492 509 BY SIMILARITY.  
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 331 331 F -> T (IN REF. 2).  
FT CONFLICT 362 362 N -> D (IN REF. 2).  
FT CONFLICT 374 374 R -> G (IN REF. 2).  
FT CONFLICT 404 404 R -> T (IN REF. 2).  
FT CONFLICT 552 552 F -> V (IN REF. 2).  
SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match 90.8%; Score 934.5; DB 1; Length 564;  
Best Local Similarity 92.0%; Pred. No. 2.3e-79;  
Matches 183; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 2 IFPKQYPIINFTAGATVQSYNFIKAVRGRLLTGADVRHEIPVLPNVRVGLPINQRFILV 61  
DB 25 IFPKQYPIINFTAGATVQSYNFIKAVRGRLLTGADVRHEIPVLPNVRVGLPINQRFILV 84

QY 62 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDQDAEAIHTLFTDVQNRVYTFAG 121  
DB 85 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDQDAEAIHTLFTDVQNRVYTFAG 144

QY 122 GNYDRLEQLAGNLENIELGNGPLEBAISALYYSTGGTQPLTARSFFIOMISEAR 181  
DB 145 GNYDRLEQLAGNLENIELGNGPLEBAISALYYSTGGTQPLTARSFFIOMISEAR 203

QY 182 FOYIEGEMTRIRYNRRA 200  
DB 204 FOYIEGEMTRIRYNRRA 222

RESULT 3  
ABRC\_ABRPR STANDARD; PRT; 562 AA.

ID\_ABRPR  
AC P28590;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Abrin-c precursor [Contains: Abrin-c A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-c B chain].  
DE Abrus precatorius (Indian licorice) (Crab's eye).  
OS



Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
eurosids I; Fabales; Fabaceae; Papilionoideae; Abrusae; Abrus.  
NCBI\_TaxID=3816;  
[1]  
SEQUENCE FROM N.A.  
TISSUE=Leaf;  
MEDLINE=91266957; PubMed=20501419;  
Wood K.A., Lord J.M., Wawrynczak E.J., Piatak M.;  
"Preproabrin: genomic cloning, characterisation and the expression of  
the A-chain in *Escherichia coli*.";  
Eur. J. Biochem. 198.723-732 (1991).  
-!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
-!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
specific adenosine on the 28S rRNA.  
-!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
-!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
-!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
-!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
EMBL; X55667; CAA39202.1; - -  
PIR; S16022; S16022.  
HSSP; P11140; LABR.  
InterPro; IPR000772; Ricin\_B\_lectin.  
InterPro; IPR001574; RIP.  
Pfam; PF00652; Ricin\_B\_lectin; 6.  
Pfam; PF00161; RIP; 1.  
PRINTS; PR00396; SHIGARICIN.  
SMART; SM00458; RICIN; 2.  
PROSITE; PS0231; RICIN B LECTIN; 2.  
PROSITE; PS00275; SHIGA\_RICIN; 1.  
Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
Glycoprotein; Lectin; Signal; Pyrrolidone carboxylic acid.  
SIGNA1 1 34  
BY SIMILARITY  
FT CHAIN 35 285 ABRIN C A CHAIN (BY SIMILARITY).  
FT PEPTIDE 286 295 LINKER PEPTIDE (BY SIMILARITY).  
FT CHAIN 296 562 ABRIN C B CHAIN (BY SIMILARITY).  
FT DOMAIN 307 434 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 437 561 RICIN B-TYPE LECTIN 2.  
FT REPEAT 317 359 1-ALPHA.  
FT REPEAT 360 400 1-BETA.  
FT REPEAT 403 435 1-GAMMA.  
FT REPEAT 448 483 2-ALPHA.  
FT REPEAT 487 526 2-BETA.  
FT REPEAT 529 562 2-GAMMA.  
FT ACT\_SITE 198 198 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 281 303 BY SIMILARITY.  
FT DISULFID 320 339 BY SIMILARITY.  
FT DISULFID 363 380 BY SIMILARITY.  
FT DISULFID 451 464 BY SIMILARITY.  
FT DISULFID 490 507 BY SIMILARITY.  
FT MOD\_RES 35 35 PYRROLIDONE CARBOXYLIC ACID  
(BY SIMILARITY).  
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SO SEQUENCE 562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;

## Query Match

```

Best Local Similarity 45.5%; Pred.No. 2.6e-25;
Matches 86; Conservative 24; Mismatches 70; Indels 9; Gaps 4;

QY      6 QYPIINFTTAGATVQSYTNFIRAVRGRLTGDVRRHEIPVLPNRVGLPINQRFILVELSN 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      35 QDOVKFTTEGATSQSKQFIEALRQLTGG--LIHDIPVLDPPTTVEERNRYITVELSN 92
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      66 HAELSVTALDVNTAVVYVGRAGNSAYFFHPDQNOEDAEATHLETFDVQNYTTFARGGNYD 125
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      93 SERSETEVGIDVTNAVYAVRAGSQSYFL---RDAPASASTYLPFGTQ-RYSLEFDSYG 148
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      126 RLRLQAGNLRENTELGNGPLFEAEISALYVYVSTGTQGLPTLARGSFIIQIMISEAARFYI 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      149 DLEWAHQFTEELSLGLOALTHAIS---FLRSGASNDEEKARTLIVIQMASEAARYXI 205
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      186 EGMRTTRIR 194
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      206 SNRVGVGSIR 214
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
ABRB ABRPR
ID ABRB ABRPR STANDARD; PRT; 527 AA.
AC Q06077; P81374;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Abrin-b precursor [Contains: Abrin-b A chain (rRNA N-glycosidase)
DE (EC 3.2.2.2); Abrin-b B chain].
OS Abrus precatorius (Indian locoweed) (Crab's eye).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
XX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93132798; PubMed=7763422;
RX Kimura M., Sumizawa T., Funatsu G.;
RT "The complete amino acid sequences of the B-chains of abrin-a and
RT abrin-b, toxic proteins from the seeds of Abrus precatorius.";
RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRIN-A IS MORE TOXIC THAN RACIN.
CC -!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC PRECEDES ENDOCYTOSIS.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----

```

## RESULT 5



CC	EMBL; AB000666; BAA21786.1; -	
DR	PIR; JC5606; JC5606.	
DR	PIR; J00393; J00393.	
DR	HSSP; P09989; ILMRJ.	
DR	InterPro; IPR001574; RIP.	
DR	Pfam; PF00161; RIP; 1.	
DR	PRINTS; PR00396; SHIGARICIN.	
DR	PROSITE; PS00275; SHIGA RICIN; 1.	
KW	Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;	
	Toxin; Signal.	
FT	SIGNAL 1 21 POTENTIAL..	
FT	CHAIN 22 270 KARASURIN-C.	
FT	CHAIN 24 270 KARASURIN-A.	
FT	PROPEP 271 289 REMOVED IN MATURE FORM.	
FT	ACT_SITE 183 183 BY SIMILARITY.	
FT	SEQUENCE 289 AA; 31704 MW; 983D3E3242887B26 CRC64;	
QY	Query Match 32.8%; Score 338; DB 1; Length 289;	
DB	Best Local Similarity 40.0%; Pred. No. 3.2e-24;	
	Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5	
QY	10 INFTTAGATVQSYTNFIRAVRGRLLTTCADVRHEIPVLNPKRVGLPINQRFILVELSNHAEL 69	
DB	25 VSRFLSGATSSYGVFISNRKALPYERKI-YDIPLL--RSTLPQSQRVALHLTNVADE 81	
QY	70 SVTLALDVNAYVGVYRAGNSAYFFHFDNQEDA-EAITHLFTDVQNKRYTFATGGNYDRLE 128	
DB	82 TTSVAIDVTNVYMGVYRAGDTSYFF--NEASATEAKYVFKDAKKVTLTPYSGNYERLQ 138	
QY	129 QLQGNRENIELNGNPLEEASIALYVYSTGGTQTLPLARSFICIQMISEAARFQVIEGE 188	
DB	139 TAAGKIRENIPGLPALDSAITLLFYNN-----SAASALMVLIQSTSEAAKYKFIQQ 193	
QY	189 METRI 193	
DB	194 IGRKV 198	
RESULT 8		
ID	NIGB SAMNI STANDARD; PRT; 563 AA.	
AC	P33183; P33184; P93542;	
DT	01-OCT-1993 (Rel. 27, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Nigrin b precursor (Agglutinin V) (SNAV) [contains: Nigrin b A chain	
DE	(rRNA N-glycosidase) (EC 3.2.2.22; Nigrin b B chain).	
OS	Sambucus nigra (European elder).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
OC	Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.	
OX	NCBI_Taxid=4202;	
RX	SEQUENCE FROM N.A.	
RC	TISSUE=Barb;	
RC	MDLINE=96215449; PubMed=8647092;	
RA	Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;	
RT	"Characterization and molecular cloning of Sambucus nigra agglutinin V	
RT	(nigrin b), a GalNAc-specific type-2 ribosome-inactivating protein	
RT	from the bark of elderberry (Sambucus nigra).";	
RL	Eur. J. Biochem. 237:505-513 (1996).	
RN	[2]	
RP	SEQUENCE OF 26-49 AND 298-321.	
RC	TISSUE=Barb;	
RX	MDLINE=94003077; PubMed=8400135;	
RA	Girbes T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R.,	
RA	Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;	
RT	"Isolation and partial characterization of nigrin b, a non-toxic	
RT	novel type 2 ribosome-inactivating protein from the bark of Sambucus	
RT	nigra L.";	
RL	Plant Mol. Biol. 22:1181-1186 (1993).	
CC	-!- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN	

CC PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN  
 CC SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B-CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 CC BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES  
 CC ENDOCYTOSIS.  
 CC  
 CC -|- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -|- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -|- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -|- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U41299; AAB39475.1; --  
 CC FIRM: S37382; S37382; Ricin\_B\_lectin.  
 CC InterPro: IPR000772; Ricin\_B\_lectin.  
 CC InterPro: IPR001574; RIP.  
 CC Pfam: PF00652; Ricin\_B\_lectin; 6.  
 CC Pfam: PF00161; RIP; 1.  
 CC PRINTS: PR00396; SHIGARICIN.  
 CC SMART: SM00458; RICIN; 2.  
 CC PROSITE: PS00275; SHIGA\_RICIN; 1.  
 CC PROSITE: PS0231; RICIN\_B\_LECTIN; 2.  
 CC KX plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 CC Glycoprotein; Lectin; Signal.  
 CC SIGNAL 1 25  
 CC CHAIN 26 297 NIGRIN B A CHAIN.  
 CC CHAIN 298 563 NIGRIN B B CHAIN.  
 CC DOMAIN 305 431 RICIN B-TYPE LECTIN 1.  
 CC DOMAIN 434 559 RICIN B-TYPE LECTIN 2.  
 CC REPEAT 316 356 1-ALPHA.  
 CC REPEAT 357 397 1-BETA.  
 CC REPEAT 400 432 1-GAMMA.  
 CC REPEAT 445 482 2-ALPHA.  
 CC REPEAT 486 524 2-BETA.  
 CC REPEAT 527 554 2-GAMMA.  
 CC ACT SITE 188 188 BY SIMILARITY.  
 CC DISULFID 274 302 INTERCHAIN (BY SIMILARITY).  
 CC DISULFID 319 338 BY SIMILARITY.  
 CC DISULFID 360 377 BY SIMILARITY.  
 CC DISULFID 448 463 BY SIMILARITY.  
 CC DISULFID 489 506 BY SIMILARITY.  
 CC CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 39 39 K -> V (IN REF. 2).  
 CC CONFLICT 39 39  
 CC SEQUENCE 563 AA; 62300 MW; F250CBE24621BF14 CRC64;  
 CC  
 CC Query Match 32.1%; Score 330.5; DB 1; Length 563;  
 CC Best Local Similarity 39.5%; Pred. No. 3.7e-23;  
 CC Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;  
 CC  
 CC 7 YPIINFTAGTAVSYNTFIRAVRGLTGTADVREHIVPLNVRVGLPINORFILVELSNH 66  
 CC 28 YPSVSNLDGAKSYRDFLSNLRKATVGTGYEVNGLPVLRRSEVQKSRFVLVPLTNY 87  
 CC 67 AELSVTLALDYNVAVGYRAGNAGYAPHPNQDAEAI--THLFTDVQNYRTAFGNY 124  
 CC 88 NGNTVTLAVDYNLYVAFSGNANSYFF-----KDTEVQKSNLFGVTKQN-TLSFTGNY 141  
 CC 125 DRLEQAGNLENGELNGPLEEISALYYSTGTQTLPLARSFIICMISEAAARFOY 184

Db 142 DNLEAANTRRRESIELGSPDIGNAITSYHGD-----SVARSLLVVIQWSEARFRY 194  
 QY 185 IEGEMRTRIR 194  
 Db 195 IEQEVRRSLQ 204  
 RESULT 9  
 ABRERA ABRPR  
 ID ABRERA ABRPR STANDARD; PRT; 528 AA.  
 AC P11140; P28589;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ABRin-a precursor [Contains: ABRin-a A chain (rRNA N-glycosidase)  
 (EC 3.2.2.22); ABRin-a B chain].  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OC NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoabrinins determined by cDNA  
 sequencing. Conservation and significance."  
 RL J. Mol. Biol. 229:263-267(1993).  
 RN [2]  
 RP SEQUENCE OF 1-251.  
 RC TISSUE=Seed;  
 RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;  
 RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic  
 protein from the seeds of Abrus precatorius."  
 RL Agric. Biol. Chem. 52:1095-1097(1988).  
 RN [3]  
 RP SEQUENCE OF 1-251 FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE=91201329; PubMed=2016300;  
 RA Evensen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin  
 A-chains."  
 RL J. Biol. Chem. 266:6848-6852(1991).  
 RN [4]  
 RP SEQUENCE OF 262-528.  
 RX MEDLINE=92371656; PubMed=1505674;  
 RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;  
 RT "The complete primary structure of abrin-a B chain."  
 RL FEBS Lett. 309:115-118(1992).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).  
 RX MEDLINE=95333188; PubMed=7608980;  
 RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;  
 RT "Crystal structure of abrin-a at 2.14 A."  
 RL J. Mol. Biol. 250:354-367(1995).  
 CC -|- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -|- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 CC PRECEDES ENDOCYTOSIS.  
 CC -|- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -|- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -|- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -|- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -|- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M98344; AAA32624.1; AUT\_INIT.  
 DR EMBL; X54872; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S32429; TZLSA.  
 DR PDB; 1ABR; 07-FEB-95.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.  
 FT CHAIN 1 251  
 FT PEPTIDE 252 261  
 FT CHAIN 262 528  
 FT DOMAIN 273 400  
 FT DOMAIN 403 527  
 FT REPEAT 283 325  
 FT REPEAT 326 366  
 FT REPEAT 369 401  
 FT REPEAT 414 449  
 FT REPEAT 453 492  
 FT REPEAT 495 528  
 FT ACT\_SITE 164 164  
 FT DISULFID 247 269  
 FT DISULFID 286 305  
 FT DISULFID 329 346  
 FT DISULFID 417 430  
 FT DISULFID 456 473  
 FT MOD\_RES 1 1  
 FT CARBOHYD 361 361  
 FT CARBOHYD 401 401  
 FT CONFLICT 202 202  
 FT CONFLICT 298 298  
 FT CONFLICT 427 427  
 FT CONFLICT 467 467  
 FT CONFLICT 483 483  
 FT STRAND 5 8  
 FT TURN 10 11  
 FT TURN 14 28  
 FT STRAND 32 33  
 FT TURN 34 35  
 FT TURN 42 43  
 FT STRAND 47 49  
 FT STRAND 51 57  
 FT STRAND 63 69  
 FT TURN 70 72  
 FT STRAND 75 79  
 FT STRAND 83 86  
 FT TURN 88 89  
 FT TURN 92 93  
 FT TURN 94 97  
 FT TURN 100 101  
 FT STRAND 103 106  
 FT TURN 113 114  
 FT TURN 115 119  
 FT TURN 124 126  
 FT STRAND 129 129  
 FT HELIX 131 142  
 FT TURN 143 144  
 FT HELIX 148 167  
 FT STRAND 168 168  
 FT HELIX 169 180  
 FT TURN 181 182

FT STRAND 185  
 FT HELIX 186  
 FT TURN 187  
 FT HELIX 188  
 FT STRAND 189  
 FT TURN 190  
 FT STRAND 191  
 FT TURN 192  
 FT STRAND 193  
 FT TURN 194  
 FT STRAND 195  
 FT TURN 196  
 FT STRAND 197  
 FT TURN 198  
 FT STRAND 199  
 FT TURN 200  
 FT STRAND 201  
 FT TURN 202  
 FT STRAND 203  
 FT TURN 204  
 FT STRAND 205  
 FT TURN 206  
 FT STRAND 207  
 FT TURN 208  
 FT STRAND 209  
 FT TURN 210  
 FT STRAND 211  
 FT TURN 212  
 FT STRAND 213  
 FT TURN 214  
 FT STRAND 215  
 FT TURN 216  
 FT STRAND 217  
 FT TURN 218  
 FT STRAND 219  
 FT TURN 220  
 FT STRAND 221  
 FT TURN 222  
 FT STRAND 223  
 FT TURN 224  
 FT STRAND 225  
 FT TURN 226  
 FT STRAND 227  
 FT TURN 228  
 FT STRAND 229  
 FT TURN 230  
 FT STRAND 231  
 FT TURN 232  
 FT STRAND 233  
 FT TURN 234  
 FT STRAND 235  
 FT TURN 236  
 FT STRAND 237  
 FT TURN 238  
 FT STRAND 239  
 FT TURN 240  
 FT STRAND 241  
 FT TURN 242  
 FT STRAND 243  
 FT TURN 244  
 FT STRAND 245  
 FT TURN 246  
 FT STRAND 247  
 FT TURN 248  
 FT STRAND 249  
 FT TURN 250  
 FT STRAND 251  
 FT TURN 252  
 FT STRAND 253  
 FT TURN 254  
 FT STRAND 255  
 FT TURN 256  
 FT STRAND 257  
 FT TURN 258  
 FT STRAND 259  
 FT TURN 260  
 FT STRAND 261  
 FT TURN 262  
 FT STRAND 263  
 FT TURN 264  
 FT STRAND 265  
 FT TURN 266  
 FT STRAND 267  
 FT TURN 268  
 FT STRAND 269  
 FT TURN 270  
 FT STRAND 271  
 FT TURN 272  
 FT STRAND 273  
 FT TURN 274  
 FT STRAND 275  
 FT TURN 276  
 FT STRAND 277  
 FT TURN 278  
 FT STRAND 279  
 FT TURN 280  
 FT STRAND 281  
 FT TURN 282  
 FT STRAND 283  
 FT TURN 284  
 FT STRAND 285  
 FT TURN 286  
 FT STRAND 287  
 FT TURN 288  
 FT STRAND 289  
 FT TURN 290  
 FT STRAND 291  
 FT TURN 292  
 FT STRAND 293  
 FT TURN 294  
 FT STRAND 295  
 FT TURN 296  
 FT STRAND 297  
 FT TURN 298  
 FT STRAND 299  
 FT TURN 300  
 FT STRAND 301  
 FT TURN 302  
 FT STRAND 303  
 FT TURN 304  
 FT STRAND 305  
 FT TURN 306  
 FT STRAND 307  
 FT TURN 308  
 FT STRAND 309  
 FT TURN 310  
 FT STRAND 311  
 FT TURN 312  
 FT STRAND 313  
 FT TURN 314  
 FT STRAND 315  
 FT TURN 316  
 FT STRAND 317  
 FT TURN 318  
 FT STRAND 319  
 FT TURN 320  
 FT STRAND 321  
 FT TURN 322  
 FT STRAND 323  
 FT TURN 324  
 FT STRAND 325  
 FT TURN 326  
 FT STRAND 327  
 FT TURN 328  
 FT STRAND 329  
 FT TURN 330  
 FT STRAND 331  
 FT TURN 332  
 FT STRAND 333  
 FT TURN 334  
 FT STRAND 335  
 FT TURN 336  
 FT STRAND 337  
 FT TURN 338  
 FT STRAND 339  
 FT TURN 340  
 FT STRAND 341  
 FT TURN 342  
 FT STRAND 343  
 FT TURN 344  
 FT STRAND 345  
 FT TURN 346  
 FT STRAND 347  
 FT TURN 348  
 FT STRAND 349  
 FT TURN 350  
 FT STRAND 351  
 FT TURN 352  
 FT STRAND 353  
 FT TURN 354  
 FT STRAND 355  
 FT TURN 356  
 FT STRAND 357  
 FT TURN 358  
 FT STRAND 359  
 FT TURN 360  
 FT STRAND 361  
 FT TURN 362  
 FT STRAND 363  
 FT TURN 364  
 FT STRAND 365  
 FT TURN 366  
 FT STRAND 367  
 FT TURN 368  
 FT STRAND 369  
 FT TURN 370  
 FT STRAND 371  
 FT TURN 372  
 FT STRAND 373  
 FT TURN 374  
 FT STRAND 375  
 FT TURN 376  
 FT STRAND 377  
 FT TURN 378  
 FT STRAND 379  
 FT TURN 380  
 FT STRAND 381  
 FT TURN 382  
 FT STRAND 383  
 FT TURN 384  
 FT STRAND 385  
 FT TURN 386  
 FT STRAND 387  
 FT TURN 388  
 FT STRAND 389  
 FT TURN 390  
 FT STRAND 391  
 FT TURN 392  
 FT STRAND 393  
 FT TURN 394  
 FT STRAND 395  
 FT TURN 396  
 FT STRAND 397  
 FT TURN 398  
 FT STRAND 399  
 FT TURN 400  
 FT STRAND 401  
 FT TURN 402  
 FT STRAND 403  
 FT TURN 404  
 FT STRAND 405  
 FT TURN 406  
 FT STRAND 407  
 FT TURN 408  
 FT STRAND 409  
 FT TURN 410  
 FT STRAND 411  
 FT TURN 412  
 FT STRAND 413  
 FT TURN 414  
 FT STRAND 415  
 FT TURN 416  
 FT STRAND 417  
 FT TURN 418  
 FT STRAND 419  
 FT TURN 420  
 FT STRAND 421  
 FT TURN 422  
 FT STRAND 423  
 FT TURN 424  
 FT STRAND 425  
 FT TURN 426  
 FT STRAND 427  
 FT TURN 428  
 FT STRAND 429  
 FT TURN 430  
 FT STRAND 431  
 FT TURN 432  
 FT STRAND 433  
 FT TURN 434  
 FT STRAND 435  
 FT TURN 436  
 FT STRAND 437  
 FT TURN 438  
 FT STRAND 439  
 FT TURN 440  
 FT STRAND 441  
 FT TURN 442  
 FT STRAND 443  
 FT TURN 444  
 FT STRAND 445  
 FT TURN 446  
 FT STRAND 447  
 FT TURN 448  
 FT STRAND 449  
 FT TURN 450  
 FT STRAND 451  
 FT TURN 452  
 FT STRAND 453  
 FT TURN 454  
 FT STRAND 455  
 FT TURN 456  
 FT STRAND 457  
 FT TURN 458  
 FT STRAND 459  
 FT TURN 460  
 FT STRAND 461  
 FT TURN 462  
 FT STRAND 463  
 FT TURN 464  
 FT STRAND 465  
 FT TURN 466  
 FT STRAND 467  
 FT TURN 468  
 FT STRAND 469  
 FT TURN 470  
 FT STRAND 471  
 FT TURN 472  
 FT STRAND 473  
 FT TURN 474  
 FT STRAND 475  
 FT TURN 476  
 FT STRAND 477  
 FT TURN 478  
 FT STRAND 479  
 FT TURN 480  
 FT STRAND 481  
 FT TURN 482  
 FT STRAND 483  
 FT TURN 484  
 FT STRAND 485  
 FT TURN 486  
 FT STRAND 487  
 FT TURN 488  
 FT STRAND 489  
 FT TURN 490  
 FT STRAND 491  
 FT TURN 492  
 FT STRAND 493  
 FT TURN 494  
 FT STRAND 495  
 FT TURN 496  
 FT STRAND 497  
 FT TURN 498  
 FT STRAND 499  
 FT TURN 500  
 FT STRAND 501

Query Match 32.0%; Score 329.5; DB 1; Length 528;  
 Best Local Similarity 42.8%; Pred. No. 4.2e-23;  
 Matches 80; Conservative 28; Mismatches 66; Indels 13; Gaps 5;

QY 10 INFATAGATVQSYNTNFTIRAVRGRLTTTCADYRHEIPVLPNRVGLPINORFILLVLSNHAEL 69  
 DB 5 IKFTEGATQSQYKQFIEALRRLRGG--LHIDIPVLDPDTTLOERNRITIVLSNDSDE 62  
 QY 70 SVTLALDVTNAYVVGVRAGNSAYEFH--PDNQEDAEAITHLFTDVQNRITFAFGNYDRL 127

Db 63 SIEGVIDTVAAYVAYRAGTQSYFLRDAPSSAD-----YLFTGT-DQHSLPFGTYGDL 116  
QY 128 BOLAGNRENLGNLPEEALISALYYSTGTQPTLARGFIICIMISEAARFOIEG 187  
Db 117 ERWQHQRQIPLGLQALTHGIS---FFRSGNDNEEKARLLVLIQWVAARFRYSN 173

QY 188 EMRTIR 194

Db 174 RRVSIQ 180

## RESULT 10

RIPI\_BRYDI  
ID RIPI\_BRYDI STANDARD; PRT; 290 AA.  
AC P33185; Q9S819;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)  
DE (EC 3.2.2.22) (BD1).  
OS Bryonia dioica (Red bryony).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
OX NCBI\_TaxID=3652;  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RC TISSUE=Leaf;  
RX MEDLINE=97228081; PubMed=9115985;  
RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,  
RA Siegall C.B.;  
RT "Molecular, biological, and preliminary structural analysis of  
RT recombinant bryodin I, a ribosome-inactivating protein from the plant  
RT Bryonia dioica.";  
RL Biochemistry 36:3095-3103(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Siegall C.B.;  
RT "Cloning and expression of a gene encoding bryodin I from Bryonia  
RT dioica.";  
RL Patent number US5541110, 30-JUL-1996.  
RN [3]  
RP SEQUENCE OF 24-66.  
RC TISSUE=Seed;  
RX MEDLINE=89326691; PubMed=2753596;  
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Sorio M.,  
RA Lappi D.;  
RT "N-terminal sequence of some ribosome-inactivating proteins.";  
RL Int. J. Pept. Protein Res. 33:263-267(1989).  
RN [4]  
RP SEQUENCE OF 24-43.  
RC TISSUE=Root;  
RX MEDLINE=95151812; PubMed=7849072;  
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
RA Marquardt H.;  
RT "Characterization of ribosome-inactivating proteins isolated from  
RT Bryonia dioica and their utility as carcinoma-reactive  
RT immunoconjugates.";  
RL Bioconj. Chem. 5:423-429(1994).  
CC -|- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
CC PROTEIN SYNTHESIS IN ANIMAL CELLS.  
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S rRNA.  
CC -|- PFM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO  
CC PRODUCE A SHORTER PROTEIN.  
CC -|- BIOTECHNOLOGY: Especially useful as immunotoxin for  
CC pharmacological applications as it has low toxicity in rats and  
CC mice but is potent once inside target cells.  
CC -|- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC TYPE 1 RIP SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; I24020; -, NOT\_ANNOTATED\_CDS.  
DR PIR; S16491; S16491.  
DR PDB; 1BRY; 04-MAR-98.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
KW 3D-structure; Multigene family; Glycoprotein; Signal.  
FT SIGNAL 1 23 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.  
FT CHAIN 24 270 MISSING IN MATURE PROTEIN.  
FT PROPEP 271 290 BY SIMILARITY.  
FT ACT\_SITE 183 183  
FT ACT\_SITE 212 212  
FT CARBOHYD 214 214  
FT CARBOHYD 250 250  
FT MUTAGEN 212 212  
FT CONFLICT 61 65 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT STRAND 25 28 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TURN 30 31 E->K: REDUCES ACTIVITY 10-FOLD.  
FT HELIX 34 46 RSSIS -> LRHXI (IN REF. 3).  
FT TURN 47 47  
FT STRAND 50 54  
FT TURN 55 56  
FT STRAND 57 60  
FT HELIX 66 69  
FT STRAND 70 76  
FT TURN 78 79  
FT STRAND 82 88  
FT TURN 89 92  
FT STRAND 93 99  
FT TURN 100 101  
FT STRAND 102 105  
FT HELIX 109 114  
FT TURN 115 117  
FT TURN 120 121  
FT STRAND 124 127  
FT HELIX 134 141  
FT TURN 142 142  
FT HELIX 145 147  
FT STRAND 150 150  
FT HELIX 152 163  
FT TURN 164 165  
FT HELIX 167 186  
FT STRAND 187 187  
FT HELIX 188 196  
FT STRAND 202 202  
FT HELIX 206 213  
FT TURN 214 214  
FT HELIX 215 225  
FT TURN 226 230  
FT STRAND 231 239  
FT TURN 241 242  
FT STRAND 245 250  
FT TURN 251 252  
FT HELIX 254 257  
FT TURN 258 259  
FT STRAND 260 260  
FT STRAND 263 263  
FT HELIX 266 268  
SQ SEQUENCE 290 AA; E965CD9C031A42DB CRC64;

Query Match 30.3%; Score 312; DB 1; Length 290;  
Best Local Similarity 37.3%; Pred. No. 8.5e-22;  
Matches 69; Conservative 46; Mismatches 58; Indels 12; Gaps 5;



QY 10 INFATTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRVGLPINQRFILVELSNHAE 69  
 DB 25 VPSRLSGATTTSYGVFIKNLREALPYERKV-YNIPLL--RSSISGSGRYTLHLHTNYADE 81  
 QY 70 SVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 DB 82 TISVADVTVNVYMGVLGVDSVFF--NEASATEAKVFKDAKKVTLPYSGNYERLUQ 138  
 QY 129 QLAGNLENIELNGPLEEAISALYYSTGGTQPLTARSFICIQMISAAARFOVIEGE 188  
 DB 139 TAAGKIRENIPGLPALDSAITLYYYTAS-----SAASALLVLIOSTABSAKYKTEQ 193  
 QY 189 MRTRI 193  
 DB 194 IGKRV 198

## RESULT 11

RIPB\_LUFCY STANDARD; PRT; 250 AA.  
 AC P22851;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein luffin-B (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 OS Luffa cylindrica (Smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 OX NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE.  
 RT TISSUE=Seed;  
 RC MEDLINE=91248488; PubMed=1368666;  
 RA Islam M.R., Hirayama H., Funatsu G.;  
 RT "Complete amino acid sequence of luffin-b, a ribosome-inactivating protein from sponge gourd (Luffa cylindrica) seeds.";  
 RL Agric. Biol. Chem. 55:229-238(1991).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC PIR: JN0108; JN0108.  
 DR HSSP; P16094; LAHC.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin.  
 FT ACT\_SITE 160 160 BY SIMILARITY.  
 SQ SEQUENCE 250 AA; 27293 MW; F01A8DC8A107800 CRC64;

Query Match 29.2%; Score 300.5; DB 1; Length 250;  
 Best Local Similarity 34.2%; Pred. No. 8.3e-21;  
 Matches 64; Conservative 48; Mismatches 64; Indels 11; Gaps 3;  
 QY 10 INFATTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRVGLPINQRFILVELSNHAE 69  
 DB 3 VPSRLSGADSKSYKSFITAKLALPKSKVSNIPLLPSASGA---SKYILMQLSNYDAK 59  
 QY 70 SVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLEQ 129  
 DB 60 AITMAIDVTNVYMGVLGVDSVFF--ANESDAKLASQVFKGSLVTLVTPYSGNYERLQN 116  
 QY 130 LAGNLENIELNGPLEEAISALYYSTGGTQPLTARSFICIQMISAAARFOVIEGEM 189  
 DB 117 AAGKIREKIPGLFRALDSALTSIFHYDS-----TAAAFVLVLTQTAESRFFKIEQI 171  
 QY 190 RTIRYN 196  
 DB 172 IERIPKN 178

## RESULT 12

RIP1\_CUCFI STANDARD; PRT; 286 AA.  
 AC Q9FRX4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 OS Cucumis figarei.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=131071;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada T., Ohki S.T., Osaki T.;  
 RT "Cloning and analysis of a cDNA coding a putative ribosome-inactivating protein from Cucumis figarei.";  
 RL Plant Biotechnol. 17:337-340(2000).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; AB045560; BAB19677.1; -.  
 DR HSSP; P16094; LAHC.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 286 PUTATIVE RIBOSOME-INACTIVATING PROTEIN.  
 FT ACT\_SITE 185 185 BY SIMILARITY.  
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 286 AA; 31771 MW; 4EFD4966E604DA41 CRC64;

Query Match 28.8%; Score 296.5; DB 1; Length 286;  
 Best Local Similarity 34.4%; Pred. No. 2.3e-20;  
 Matches 65; Conservative 47; Mismatches 64; Indels 13; Gaps 5;  
 QY 10 INFATTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRVGLPINQRFILVELSNHAE 68  
 DB 28 VKFSLGSDSKSYKSFITSMNALPNAGDI-YNIPLLVPSISG---SRVILMQLSNYVEG 83  
 QY 69 LSVTLALDVTNAVYVGVYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 DB 84 NTITMAIDVTNVYMGVLGVDSVFF---NETDAQLASKFVFGTKSITLPYSNGYQKLQ 140  
 QY 129 QLAGNLENIELNGPLEEAISALYYSTGGTQPLTARSFICIQMISAAARFOVIEGE 188  
 DB 141 SVAEKREKIPGLFRALDSALTSIYYDSRAPI-----AFLVLIQTAAEARVYKIEQ 195  
 QY 189 MRTRIRYN 197  
 DB 196 IIDRISVSK 204

## RESULT 13

MLA\_VISAL ID MLA\_VISAL STANDARD; PRT; 254 AA.

AC P81446;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-glycosidase) (EC 3.2.2.22)  
 DE Viscum album (European mistletoe)  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Santalales; Santalaceae; Viscum.  
 OX NCBI\_TaxID=3972;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-Subsp. album;  
 RX MEDLINE=97134581; PubMed=8980141;  
 RA Soler M.H., Stoeva S., Schwaborn C., Wilhelm S., Stiefel T.,  
 RA Voelter W.;  
 RT "Complete amino acid sequence of the A chain of mistletoe lectin I.";  
 RL FEBS Lett. 399:153-157(1996).  
 CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY  
 CC INTO THE CELL OF THE A CHAIN; B CHAINS ARE ALSO RESPONSIBLE FOR  
 CC CELL AGGLUTINATION (LECTIN ACTIVITY).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 CC -!- PHARMACEUTICAL: Due to its immunomodulative effects it is being  
 CC studied in clinical trials in cancer patients as it may slow the  
 CC growth of cancer cells and be an effective treatment for solid  
 CC tumors.  
 CC -!- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA  
 CC AND NON-GLYCOSYLATED FORM MLA'.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 2 RIP SUBFAMILY.  
 DR PIR; PD0018; PD0018.  
 DR HSP; P11140; LABR.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; Repeat; Glycoprotein; Lectin.  
 KW Plant defense; Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.  
 FT ACT SITE 165  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .).  
 FT VARIANT 15 15 E -> D (IN MLA').  
 FT VARIANT 66 66 V -> I (IN MLA').  
 FT VARIANT 112 112 N -> T (IN MLA').  
 FT VARIANT 116 116 P -> T (IN MLA').  
 FT VARIANT 133 134 DQ -> EE (IN MLA').  
 FT VARIANT 140 140 T -> S (IN MLA').  
 FT VARIANT 144 144 F -> Y (IN MLA').  
 FT VARIANT 151 151 T -> A (IN MLA').  
 FT VARIANT 151 151 Y -> D (IN MLA').  
 FT VARIANT 179 179 A -> E (IN MLA').  
 FT VARIANT 184 184 V -> M (IN MLA').  
 FT VARIANT 190 190 I -> F (IN MLA').  
 FT VARIANT 218 218 PP -> ST (IN MLA').  
 FT VARIANT 223 224 T -> S (IN MLA').  
 FT VARIANT 231 231 D -> S (IN MLA').  
 FT VARIANT 235 235 D -> S (IN MLA').  
 SQ SEQUENCE 254 AA; 28478 MW; 53BAF98D3E0FF67 CRC64;  
 Query Match 27.8%; Score 286; DB 1; Length 254;  
 Best Local Similarity 38.4%; Pred. No. 1.9e-19;  
 Matches 81; Conservative 36; Mismatches 56; Indels 38; Gaps 10;  
 QY 14 TAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRRVGLPINCORFILVELSHNELSV 71  
 Db 9 THQTGGEYFRFTLLRDVSSGS-FSNEIPLL-RQSTIPVSDAQRFLVELTNGQGDV 66  
 QY 72 TLADVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQNYTFAFGGNYDLRLQLA 131

Db 67 TAAIDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQNYTFAFGGNYDLRLQLA 122  
 QY 132 GNLRNLELGNGLPEEAISALYYSTGTGTOPTLARSFFIIICMISEAARF-----Q 183  
 Db 123 GH-RDQPLGIDQLQSVTALRF---PGSRTQARSILILQIMISEAARFNPLWRYRQ 178  
 QY 184 YIE-----GEMRTIRYN 196  
 Db 179 YINSGASFLPDVYMLETSGWQOSTQVQHS 209  
 RESULT 14  
 ID RIPA\_LUFCY STANDARD; PRT; 277 AA.  
 AC Q00465;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein luffin-alpha precursor (rRNA  
 DE N-glycosidase) (EC 3.2.2.22).  
 OS Luffa cylindrica (Smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 OX NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed;  
 RX MEDLINE=9228316; PubMed=1600156;  
 RA Kataoka J., Habuka N., Miyano M., Masuta C., Koizumi A.;  
 RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-  
 RT inactivating protein from Luffa cylindrica.";  
 RL Plant Mol. Biol. 18:1199-1202(1992).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL; X62371; CAA44229.1; -.  
 DR PIR; S22494; S22494.  
 DR HSP; P16094; LAHC.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 FT SIGNAL 1 19 RIBOSOME-INACTIVATING PROTEIN LUFFIN-  
 FT CHAIN 20 277 ALPHA.  
 FT ACT SITE 179 179 BY SIMILARITY.  
 SQ SEQUENCE 277 AA; 30212 MW; EAI7FC27998C25AC CRC64;  
 Query Match 27.6%; Score 284.5; DB 1; Length 277;  
 Best Local Similarity 33.5%; Pred. No. 2.9e-19;  
 Matches 63; Conservative 45; Mismatches 69; Indels 11; Gaps 3;  
 QY 10 INFTTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRRVGLPINCORFILVELSHNEL 69  
 Db 22 VRESLGSSTSSYKXKIGDLRKALPSNGVYVITLLSSASGA---SRVTLMTLSNVDGK 78  
 QY 70 SVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQNYTFAFGGNYDLRLQLA 129  
 Db 79 AITVAVDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQNYTFAFGGNYDLRLQLA 135

QY 130 LAGNLRNTELGNGPLEEALISALYYSTGCTQTLPTLARSFIIQIMISEARFQYIEGEM 189  
DB 136 AGKIREKIPLGFPALDSAITTLFHYDS-----TAAAAAFLVLIQTAAEAARFKYIERHV 192  
QY 190 RTRIRYNR 197  
DB 191 IERISKNQ 198

## RESULT 15

RIP2 MOMB  
ID RIP2 MOMB STANDARD; PRT; 286 AA.  
AC P29339;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ribosome-inactivating protein momordin II precursor (xRNA  
N-glycosidase) (EC 3.2.2.22).  
OC Momordica balsamina (Bitter melon) (Balsam pear).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.  
OX NCBI\_TaxID=3672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seed;  
RX MEDLINE=93027170; PubMed=1408771;  
RA Ortigas M., Better M.;  
RT "Momordin II, a ribosome inactivating protein from Momordica  
balsamina, is homologous to other plant proteins.";  
RL Nucleic Acids Res. 20:4662-4662(1992).  
CC -|- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one  
specific adenosine on the 28S rRNA.  
CC -|- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC TYPE 1 RIP SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z12175; CAA78166.1; -;  
DR PIR; S25560; S25560.  
DR PDB; 1CF5; 07-JUN-99.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
KW 3D-structure.  
FT SIGNAL 1 23  
FT CHAIN 24 286 RIBOSOME-INACTIVATING PROTEIN MOMORDIN  
FT II.  
FT ACT SITE 181 181 BY SIMILARITY.  
FT SEQUENCE 286 AA; 32031 MW; 3B89FFFAE6B25986 CRC64;  
SQ

Query Match 27.1%; Score 279; DB 1; Length 286;  
Best Local Similarity 35.3%; Pred. No. 9.8e-19;  
Matches 66; Conservative 42; Mismatches 67; Indels 12; Gaps 5;  
QY 10 INFTTAGATVQSYTNFRAVRGRLTGADVRHEIPVLPNRVGLPQNRPILVELSNHAEL 69  
DB 25 VNFEDLSTAKTYTKFIEDFRATLPESHKV-YDIPLLYSTIS--DSRRFILLDLTSYAYE 81  
QY 70 SVTLALDVNAYVYGRAGNSAYFFHPDQDEABATHLFTDVQNRYPFAFGNYDRLEQ 129  
DB 82 TTSVAIDVTNVYVAYTRDVSYFF---KESPEAYNILFKGTR-KITLPYTGNYENLOT 137  
QY 130 LAGNLRNTELGNGPLEEALISALYYSTGCTQTLPTLARSFIIQIMISEARFQYIEGEM 189

DB 138 AAHKIRENIDLGLPALSSAITTLFYNA-----QSAPSALLVLIQTAAEAARFKYIERHV 192  
QY 190 RTRIRYN 196  
DB 193 AKYVATN 199

Search completed: February 10, 2004, 16:23:26  
Job time : 7.67161 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 26.6864 Seconds  
(without alignments)  
1933.961 Million cell updates/sec

Title: US-10-083-336A-10  
Perfect score: 1029  
Sequence: 1 MIFPKQYPIINFATTAGTQV.....RFQYIEGEMRTIRYNRSEA 200

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 23.\*

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1021	99.2	541	Q41174	Q41174 ricinus com
2	401.5	39.0	580	Q94BW3	Q94BW3 cinnamomum
3	397.5	38.6	580	Q94BW4	Q94BW4 cinnamomum
4	397.5	38.6	581	Q94BW5	Q94BW5 cinnamomum
5	395.5	38.4	549	Q9FV22	Q9FV22 cinnamomum
6	350.5	34.1	563	Q04367	Q04367 sambucus ni
7	347.5	33.8	564	Q9AVR2	Q9AVR2 sambucus eb
8	344.5	33.5	528	Q06076	Q06076 abrus prec
9	340	33.0	289	Q94KE4	Q94KE4 trichosanth
10	338	32.8	247	Q9LRE3	Q9LRE3 trichosanth
11	338	32.8	289	Q41216	Q41216 trichosanth
12	336.5	32.7	252	Q38760	Q38760 abrus prec
13	330.5	32.1	563	Q94582	Q94582 sambucus ni
14	330.5	32.1	563	Q0GT32	Q0GT32 sambucus ni
15	329.5	32.0	252	Q38761	Q38761 abrus prec
16	324	31.5	270	Q8LPV7	Q8LPV7 trichosanth

17	323.5	31.4	251	10	Q96236	Q96236 abrus prec
18	322.5	31.3	251	10	Q96237	Q96237 abrus prec
19	319	31.0	565	10	Q04071	Q04071 sambucus ni
20	317	30.8	270	10	Q41611	Q41611 trichosanth
21	316.5	30.8	251	10	Q96235	Q96235 abrus prec
22	314.5	30.6	547	10	Q9M6E9	Q9M6E9 abrus prec
23	310.5	30.2	566	10	Q04072	Q04072 sambucus ni
24	307.5	29.9	278	10	Q00980	Q00980 luffa cylin
25	291	28.3	570	10	Q41358	Q41358 sambucus ni
26	285	27.7	249	10	Q8LKQ5	Q8LKQ5 viscum albu
27	285	27.7	570	10	Q22415	Q22415 sambucus ni
28	283.5	27.6	604	10	Q9M654	Q9M654 polygonatum
29	283	27.5	251	10	Q8LKQ4	Q8LKQ4 viscum albu
30	282	27.4	254	10	Q8LKQ6	Q8LKQ6 viscum albu
31	280	27.2	264	10	Q9FSH2	Q9FSH2 momordica c
32	278	27.0	293	10	Q8S452	Q8S452 jatropa cu
33	277.5	27.0	565	10	Q8W243	Q8W243 viscum albu
34	277	26.9	286	10	Q9FUV7	Q9FUV7 momordica c
35	275	26.9	569	10	P93543	P93543 sambucus ni
36	275	26.7	531	10	Q8RXH6	Q8RXH6 viscum albu
37	273	26.5	249	10	Q8RXH7	Q8RXH7 viscum albu
38	272	26.4	286	10	Q41257	Q41257 momordica c
39	272	26.4	298	10	O04358	O04358 iris hollan
40	270	26.2	592	10	Q8W2E7	Q8W2E7 iris hollan
41	268	26.0	300	10	O04356	O04356 iris hollan
42	266.5	25.9	603	10	Q9M653	Q9M653 polygonatum
43	262.5	25.5	299	10	Q8GZN9	Q8GZN9 euphorbia s
44	261	25.4	293	10	Q8VYU0	Q8VYU0 jatropa cu
45	259	25.2	573	10	Q8W2E8	Q8W2E8 iris hollan

## ALIGNMENTS

## RESULT 1

Q41174 ID Q41174 PRELIMINARY; PRT; 541 AA.  
AC Q41174;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)  
DE Proridin A chain (EC 3.2.2.22) (rRNA N-glycosidase)  
DE (fragment).  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92338377; PubMed=1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT "Molecular cloning of ricin".  
RL Targeted Diagn. Ther. 7:81-97(1992).  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; S40366; AB22582.1; -.  
DR HSSP; F02879; IBER6.  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR InterPro; IPR001400; Somatotropin.  
DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00356; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS50231; RICIN B LECTIN; 2.  
DR PROSITE; PS00275; SHIGA RICIN; 1.  
DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
DR Hydrolase; Toxin..  
KW NON\_TER  
FT NON\_TER  
SQ SEQUENCE 541 AA; 60281 MW; 257B2CDEF1F2E9D9 CRC64;

Query Match 99.2%; Score 1021; DB 10; Length 541;  
 Best Local Similarity 99.5%; Pred. No. 9.3e-88;  
 Matches 198; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFFKQYPIINFTAGATVQSYNFIKAVRGRLTTGADVREHETPVLNVRGLPINQRFILV 61  
 DB 1 IFFKQYPIINFTAGATVQSYNFIKAVRGRLTTGADVREHETPVLNVRGLPINQRFILV 60

QY 62 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 121  
 DB 61 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120

QY 122 GNYDRLEQLAGNLRENIEGNGLPBEAISALYYSTGGTQLPFLARSFFIICIMISEAR 181  
 DB 121 GNYDRLEQLAGNLRENIEGNGLPBEAISALYYSTGGTQLPFLARSFFIICIMISEAR 180

QY 182 FOYIEGEMTRIRYNRRA 200  
 DB 181 FOYIEGEMTRIRYNRRA 199

## RESULT 2

QY4BW3 ID Q94BW3 PRELIMINARY; PRT; 580 AA.

AC Q94BW3; (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (T-EMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin III precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AY039803; AAK82460.1; -  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR Pfam; PF00161; Ricin\_B\_lectin; 6.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN III.  
 SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7F558 CRC64;

Query Match 39.0%; Score 401.5; DB 10; Length 580;  
 Best Local Similarity 50.0%; Pred. No. 3e-29;  
 Matches 94; Conservative 28; Mismatches 59; Indels 7; Gaps 5;

QY 7 YPIINFTAGATVQSYNFIKAVRGRLTTGADVREHETPVLNVRGLPINQRFILVLSN- 65  
 DB 33 YQTVTFTTKNATKTSYQFIEALRAQLASGEE-PHGIPVWRDSTVPDSKRFILVLSNW 91

QY 66 HAELSVTALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFGNYD 125  
 DB 92 AADSPVALAVDVTNAYVAYRTGSGSFLENDPDP--PAENLLPDTK-RYTFPFGSGYT 148

QY 126 RLEQLAGNLRENIEGNGLPBEAISALYYSTGGTQLPFLARSFFIICIMISEARFYI 185  
 DB 149 DLERVAGERREILLGMDFLENALSALN--NQQRALARSLLVVIQWAEAVRFRFI 206

QY 186 EGEWRTRI 193  
 DB 207 EYRVRESI 214

## RESULT 3

QY4BW4 ID Q94BW4 PRELIMINARY; PRT; 580 AA.

AC Q94BW4; (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin II precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AY039802; AAK82459.1; -  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN II.  
 SQ SEQUENCE 580 AA; 64265 MW; 37B4289ECCE0CBFF CRC64;

Query Match 38.6%; Score 397.5; DB 10; Length 580;  
 Best Local Similarity 49.5%; Pred. No. 7.1e-29;  
 Matches 93; Conservative 30; Mismatches 58; Indels 7; Gaps 5;

QY 7 YPIINFTAGATVQSYNFIKAVRGRLTTGADVREHETPVLNVRGLPINQRFILVLSN- 65  
 DB 33 YQTVTFTTKNATKTSYQFIEALRAQLASGEE-PHGIPVWRDSTVPDSKRFILVLSNW 91

QY 66 HAELSVTALDVTNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFGNYD 125  
 DB 92 AADSPVALAVDVTNAYVAYRTGSGSFLENDPDP--PAENLLPDTK-RYTFPFGSGYT 148

QY 126 RLEQLAGNLRENIEGNGLPBEAISALYYSTGGTQLPFLARSFFIICIMISEARFYI 185  
 DB 149 DLERVAGERREILLGMDFLENALSAL--WTSNINQQRALARSLLVVIQWAEAVRFRFI 206

## RESULT 4

QY4BW5 ID Q94BW5 PRELIMINARY; PRT; 581 AA.

AC Q94BW5; (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin I precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).

OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AY039801; AA82458.1; -  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 5.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS02031; RICIN\_B\_LECTIN; 2.  
 DR Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 581  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN I.  
 FT SEQUENCE 581 AA; 64215 MW; 6885F5B8FBA3D196 CRC64;  
 Query Match 38.6%; Score 397.5; DB 10; Length 581;  
 Best Local Similarity 50.0%; Pred. No. 7.1e-29;  
 Matches 94; Conservative 27; Mismatches 60; Indels 7; Gaps 5;  
 QY 7 YPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPVLPNVRGLPINORFILVELSN- 65  
 DB 33 YQVTFITTKAKTSYQTFEALRAQLASGEE-PHGIPVWRERSTVPSKRFILVELSNW 91  
 QY 66 HAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVNRVTFAPGNGYD 125  
 DB 92 AADSPVTLAVDTNAYVAVYRTGSGFFLRDNDP--PAIENLLPDTK-RYTFPFSGSYT 148  
 QY 126 RLEQLAGNLRNIEINGPLERASISALYYSTGCTGTLPTLARSFIICMISEARFOYI 185  
 DB 149 DLEGVAGRRERIEILLGMDPLENAISALWISNL--NQORALARSLLVIQMVAVRFRFI 206  
 QY 186 EGEWRTRI 193  
 DB 207 EYVRGSI 214  
 RESULT 5  
 Q9FV22 PRELIMINARY; PRT; 549 AA.  
 AC Q9FV22  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Type II ribosome-inactivating protein cinnamomin (BC 3.2.2.22) (rRNA  
 DE N-glycosidase) (fragment).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Xie L., Liu W.-Y., Wang E.-D.;  
 RT "Molecular cloning of cinnamomin A-, B-chain and the expression,  
 RT purification, characterization and mutagenesis of the A-chain";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AF259548; AAF68978.2; -  
 DR HSSP; P02879; 2AAL.

DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 5.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS02031; RICIN\_B\_LECTIN; 2.  
 DR Hydroxylase; Toxin.  
 FT NON\_TER 1  
 FT SEQUENCE 549 AA; 60648 MW; 02607FE607CA44B0 CRC64;  
 Query Match 38.4%; Score 395.5; DB 10; Length 549;  
 Best Local Similarity 50.0%; Pred. No. 1e-28;  
 Matches 94; Conservative 27; Mismatches 60; Indels 7; Gaps 5;  
 QY 7 YPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPVLPNVRGLPINORFILVELSN- 65  
 DB 1 YQVTFITTKAKTSYQTFEALRAQLASGEE-PHGIPVWRERSTVPSKRFILVELSNW 59  
 QY 66 HAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVNRVTFAPGNGYD 125  
 DB 60 AADSPVTLAVDTNAYVAVYRTGSGFFLRDNDP--PAIENLLPDTK-RYTFPFSGSYT 116  
 QY 126 RLEQLAGNLRNIEINGPLERASISALYYSTGCTGTLPTLARSFIICMISEARFOYI 185  
 DB 117 DLEGVAGRRERIEILLGMDPLENAISALWISNL--NQORALARSLLVIQMVAVRFRFI 174  
 QY 186 EGEWRTRI 193  
 DB 175 EYVRGSI 182  
 RESULT 6  
 O04367 PRELIMINARY; PRT; 563 AA.  
 ID O04367  
 AC O04367  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Ribosome inactivating protein precursor (EC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98112023; PubMed=9450339;  
 RA Van Damme E.J., Roy S., Barre A., Rouge P., Van Leuven F.,  
 RA Peumans W.J.;  
 RT "The major elderberry (Sambucus nigra) fruit protein is a lectin  
 RT derived from a truncated type 2 ribosome-inactivating protein";  
 RL Plant J. 12:1251-1260(1997).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; U76524; AAC15886.1; -  
 DR HSSP; P02879; 2AAL.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS02031; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS02075; SHIGA\_RICIN; 1.  
 KW Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 25  
 FT CHAIN 26 297  
 FT CHAIN 298 563  
 FT SEQUENCE 563 AA; 62336 MW; 3ED2B6C08E796205 CRC64;  
 SQ

Query Match 34.1%; Score 350.5; DB 10; Length 563;  
 Best Local Similarity 41.6%; Pred. No. 1.9e-24; Indels 15; Gaps 4;  
 Matches 79; Conservative 40; Mismatches 56; Indels 15; Gaps 4;

QY 7 YPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVELSNH 66  
 DB 28 YFSVSNLAGAKSATYRDFLKNLRITVAITGTYEVNGLPVLRESEVQVKNRFLVLLTNY 87

QY 67 AELSVTLALDVNAYVGVYRAGNSAYFFHPDQDEAEI--THLFTDVQNRVYTFAGGNY 124  
 DB 88 NGDVTSLAVDVTNLVYAFSANGNSYFF-----KDATLQKSNLFGVTR-QHTLPFTGNY 141

QY 125 DRLEQLAGNLRENIELGNGPLPEEAISALYYSTGGTQLPRLARSFIIICIMISEARFOY 184  
 DB 142 DNLETAAGTRRESIELGNGPLDGAITSWY--DGS-----VARSLLVLQMVPEARFRY 194

QY 185 IEGEMRTIR 194  
 DB 195 IEQEVRSLSQ 204

RESULT 7  
 Q9AVR2 PRELIMINARY; PRT; 564 AA.

AC Q9AVR2;  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein precursor (EC 3.2.2.22) (rRNA N-glycosidase).  
 DE N-glycosidase).  
 GN EBU1.  
 OS Sambucus ebulus.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=28503;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE=Leaf;  
 RA Girbes Y., Iglesias R., Perez Y., Ferreras J.M., Citores L.;  
 RT "Molecular cloning of ebulin 1";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL; A400822; CAC33178.1; -;  
 DR HSP; P02879; 2AAL.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Glycosidase; Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 298 EBU1IN L A-CHAIN.  
 FT CHAIN 299 564 EBU1IN L B-CHAIN.  
 SQ SEQUENCE 564 AA; 62694 MW; 8261681A6DB55C88 CRC64;

Query Match 33.8%; Score 347.5; DB 10; Length 564;  
 Best Local Similarity 41.1%; Pred. No. 3.6e-24;  
 Matches 78; Conservative 41; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVELSNH 66  
 DB 28 YFSVSNLAGAKSATYRDFLKNLRITVAITGTYEVNGLPVLRESEVQVKNRFLVLLTNY 87

QY 67 AELSVTLALDVNAYVGVYRAGNSAYFFHPDQDEAEI--THLFTDVQNRVYTFAGGNY 124  
 DB 88 NGDVTSLAVDVTNLVYAFSANGNSYFF-----KDATLQKSNLFLGT-TQHTLSFTGNY 141

QY 125 DRLEQLAGNLRENIELGNGPLPEEAISALYYSTGGTQLPRLARSFIIICIMISEARFOY 184  
 DB 142 DNLETAAGTRRESIELGNGPLDGAITSWY--DGS-----VARSLLVLQMVPEARFRY 194

QY 185 IEGEMRTIR 194  
 DB 195 IEQEVRSLSQ 204

RESULT 8  
 Q06076 PRELIMINARY; PRT; 528 AA.

AC Q06076;  
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE Abrin-d (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoforms determined by cDNA  
 RT sequencing: conservation and significance";  
 RL J. Mol. Biol. 229:263-267 (1993).  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL; M98346; AAA32626.1; -;  
 DR HSP; P11140; 1AUR.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 FT NON TER 1 1  
 FT NON TER 528 528  
 SQ SEQUENCE 528 AA; 58870 MW; 62ED42FB8FFE60F8 CRC64;

Query Match 33.5%; Score 344.5; DB 10; Length 528;  
 Best Local Similarity 45.0%; Pred. No. 6.3e-24;  
 Matches 85; Conservative 24; Mismatches 71; Indels 9; Gaps 4;

QY 6 QYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVELSN 65  
 DB 1 QDQVINKFTTEGATSQYKQFIEALRQLTGG--LIHDIPLVDPDPTTVEERNRYITVELSN 58

QY 66 HAELSVTLALDVNAYVGVYRAGNSAYFFHPDQDEAEI--THLFTDVQNRVYTFAGGNY 125  
 DB 59 SERESIEVGIDVTNAYVAYRAGSQSYFL---RDAPASASTYLPFGTQ-RYSLRFDGSGY 114

QY 126 RLLEQLAGNLRENIELGNGPLPEEAISALYYSTGGTQLPRLARSFIIICIMISEARFOY 185  
 DB 115 DLERWAHQTRRESIELGNGPLDGAITSWY--DGS-----VARSLLVLQMVPEARFRY 171

QY 186 EGMRTIR 194  
 DB 172 SNRVGSIR 180

RESULT 9  
 Q94KE4 PRELIMINARY; PRT; 289 AA.

ID Q94KE4  
 AC Q94KE4;  
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)



01-DEC-2001 (TrEMBLrel. 19; Last sequence update)  
01-MAR-2003 (TrEMBLrel. 23; Last annotation update)  
Trichosanthin precursor (EC 3.2.2.22) (rRNA N-glycosidase).  
TCS.  
OS Trichosanthin kirilowii (Mongolian snake-gourd).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
NCBI\_TaxID=3677;  
[1]  
SEQUENCE FROM N.A.  
Yuan H., Wang L., Wang Y., An C., Chen Z.;  
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RL CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC -1- SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AF367252; AAK52960.1; -.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Signal; Toxin.  
FT SIGNAL 1 23  
FT CHAIN 24 270  
FT SEQUENCE 289 AA; 31706 MW; AGD5602549CA5657 CRC64;  
Query Match 33.0%; Score 340; DB 10; Length 289;  
Best Local Similarity 39.5%; Pred. No. 7.2e-24;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
QY 10 INFTAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPNVRVGLPINORFILVELSNHAEL 69  
DB 25 VSRFLSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSLPGSQRYALIHLTNYADE 81  
QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
DB 82 TISVAIDVTNIVMGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPLVSGNYERLQ 138  
QY 129 QLAGNLRNIELGNPLEEASALYYSTGGTQTLPLARSFIICMISEARFOYIEGE 188  
DB 139 TAAGKIRENIPGLPALDLSAITTLFYNNAN-----SAASALMVLIOSTSEARFYIEQQ 193  
QY 189 METRI 193  
DB 194 IGRKV 198  
Query Match 33.0%; Score 340; DB 10; Length 289;  
Best Local Similarity 39.5%; Pred. No. 7.2e-24;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
QY 10 INFTAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPNVRVGLPINORFILVELSNHAEL 69  
DB 25 VSRFLSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSLPGSQRYALIHLTNYADE 81  
QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
DB 82 TISVAIDVTNIVMGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPLVSGNYERLQ 138  
QY 129 QLAGNLRNIELGNPLEEASALYYSTGGTQTLPLARSFIICMISEARFOYIEGE 188  
DB 139 TAAGKIRENIPGLPALDLSAITTLFYNNAN-----SAASALMVLIOSTSEARFYIEQQ 193  
QY 189 METRI 193  
DB 194 IGRKV 198  
RESULT 10  
Q9LRE3 ID Q9LRE3 PRELIMINARY; PRT; 247 AA.  
AC Q9LRE3  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
GN TBK.  
OS Trichosanthin sp. Bac Kan 8-98.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
NCBI\_TaxID=118182;  
[1]  
SEQUENCE FROM N.A.  
Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;  
RA "Genomic DNA clone for mature typ-1 ribosome-inactivating protein from  
RT Trichosanthin sp. sample 01 Bac Kan 8-98 Vien CINH (Hanoi).";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AB039324; BAA92530.1; -.  
DR HSSP; P09989; IMRJ.  
DR InterPro; IPR001574; RIP.

DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Toxin.  
FT NON\_TER 1 247  
FT NON\_TER 247  
FT SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;  
Query Match 32.8%; Score 338; DB 10; Length 247;  
Best Local Similarity 40.0%; Pred. No. 8.9e-24;  
Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5;  
QY 10 INFTAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPNVRVGLPINORFILVELSNHAEL 69  
DB 2 VSRFLSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSLPGSQRYALIHLTNYADE 58  
QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
DB 59 TISVAIDVTNIVMGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPLVSGNYERLQ 115  
QY 129 QLAGNLRNIELGNPLEEASALYYSTGGTQTLPLARSFIICMISEARFOYIEGE 188  
DB 116 IAAKIRENIPGLPALDLSAITTLFYNNAN-----SAASALMVLIOSTSEARFYIEQQ 170  
QY 189 METRI 193  
DB 171 IGRKV 175  
RESULT 11  
Q41216 ID Q41216 PRELIMINARY; PRT; 289 AA.  
AC Q41216  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).  
GN TRICHOSANTHIN, TCS.  
OS Trichosanthin kirilowii (Mongolian snake-gourd).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
NCBI\_TaxID=3677;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=94271613; PubMed=8003348;  
RA Zheng H., Wang B., Shaw P., Yeung H.;  
RT "[Cloning and DNA sequencing of the gene encoding trichosanthin].";  
RL I Chuan Haueh Pac 21:42-51(1994).  
CC -1- CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; S70176; AB31048.1; -.  
DR HSSP; P09989; IMRJ.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Toxin.  
FT SEQUENCE 289 AA; 31650 MW; 286AC14D48BCA175 CRC64;  
Query Match 32.8%; Score 338; DB 10; Length 289;  
Best Local Similarity 39.5%; Pred. No. 1.1e-23;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
QY 10 INFTAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPNVRVGLPINORFILVELSNHAEL 69  
DB 25 VSRFLSGATSSSYGVFISNLRKALPNERKL-YDIPLL--RSLPGSQRYALIHLTNYADE 81  
QY 70 SVTLALDVTNAYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
DB 82 TISVAIDVTNIVMGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPLVSGNYERLQ 138

QY 129 QIAGNLRNIELNGPLBEAISALYYSTGGTQLPLARSFFICIMISEARFOYIEGE 188  
Db 139 TAAGKIRENIPGLPALDSAITTFYNNAN-----SAASALMVLIOSTSEARFYKFIQQ 193  
QY 189 MKTRI 193  
Db 194 IGRV 198  
RESULT 12  
Q38760 PRELIMINARY; PRT; 252 AA.  
AC Q38760;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE ABRIN-E (RRNA N-glycosidase) (EC 3.2.2.22) (Fragment).  
GN RIP.  
OS Abrus precatorius (Indian licorice) (Crab's eye).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
OX NCBI\_TaxID=3816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LEAF;  
RX MEDLINE=91201329; PubMed=2016300;  
RA Evensen G.; Mathiesen A.; Sundan A.;  
RT "Direct molecular cloning and expression of two distinct abrin A-  
chains";  
RL J. Biol. Chem. 266:6848-6852 (1991).  
CC -1- FUNCTION. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA.  
CC ABRIN-A IS MORE TOXIC THAN RICIN.  
CC -1- FUNCTION. THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
PRECEDES ENDOCYTOSIS.  
CC -1- CATALYTIC ACTIVITY. ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -1- SUBUNIT. DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
CC -1- DOMAIN. THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
CC -1- SIMILARITY. TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING  
PROTEINS. BELONGS TO TYPE 2 RIP.  
DR EMBL; X54872; CAA38654.1; -;  
DR HSP; P11140; IABR.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.  
FT CHAIN 1 252 ABRIN E, A CHAIN (BY SIMILARITY).  
FT NON TER 252  
SQ SEQUENCE 252 AA; 28309 MW; BBFC846B9E92B5DE CRC64;  
Query Match 32.7%; Score 336.5; DB 10; Length 252;  
Best Local Similarity 44.9%; Pred. No. 1.3e-23;  
Matches 83; Conservative 24; Mismatches 69; Indels 9; Gaps 4;  
QY 10 INFETAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPENRVGLPINQRFILVELSNHAE 69  
Db 6 IKFSTEGATSQYKQFIEALRERLGG--LIHDPVLRDPTTVEERNYITVELSNSERE 63  
QY 70 SVTLALDVNTNAYVYGRAGNSAYFFHPDQDEAEAI--THLFTDVQNRYYTFAFGGNYDRLEQ 129  
Db 64 SIEVGIDVTNAYVAYRAGSOSYFL---RDAPASASTYLTGTQ--RYSLRFDGSGYGLER 119  
QY 130 LAGNLRNIELNGPLBEAISALYYSTGGTQLPLARSFFICIMISEARFOYIEGEM 189  
Db 120 WAHQTRQLSLGLOALTTAIS---FLRGSANDEKARTILVIOGMASEAARYIYISNRV 176  
QY 190 RTRIR 194

Db 177 GVSIR 181  
Q945S2 PRELIMINARY; PRT; 563 AA.  
AC Q945S2;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
N-glycosidase).  
GN AVL  
OS Sambucus nigra (European elder).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
OX NCBI\_TaxID=4202;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leaf;  
RA Van Damme E.J.M.;  
RT "Characterization and cloning of lectins and ribosome-inactivating  
proteins from Sambucus nigra leaves";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY. ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -1- SIMILARITY. BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AF409135; AAL04123.1; -;  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Toxin.  
SQ SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;  
Query Match 32.1%; Score 330.5; DB 10; Length 563;  
Best Local Similarity 39.5%; Pred. No. 1.4e-22;  
Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;  
QY 7 YPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPENRVGLPINQRFILVELSNH 66  
Db 28 YPSVSFNLDGAKSATYRDFLSNLRKTVATGTYEVNGLPVLRESEYQVKSRFVLVPLTNY 87  
QY 67 AELSVTLALDVNTNAYVYGRAGNSAYFFHPDQDEAEAI--THLFTDVQNRYYTFAFGGNY 124  
Db 88 NGMTVTILAVDVTNLYVVAEFGNSANSYFF-----KDATEVQKSNLFVGTQKN--TISFTGNY 141  
QY 125 DRLEQLAGNLRNIELNGPLBEAISALYYSTGGTQLPLARSFFICIMISEARFOY 184  
Db 142 DNLTEAANTRESIELGSPDLGATISLVHGD-----SVARSLVVIQWSEARFRY 194  
QY 185 IEQEMRTRIR 194  
Db 195 IEQEVRSIQ 204  
RESULT 14  
Q8GT32 PRELIMINARY; PRT; 563 AA.  
AC Q8GT32;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Type 2 ribosome-inactivating protein nigrin 1 precursor  
(EC 3.2.2.22).  
GN Sambucus nigra (European elder).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:20 ; Search time 32.3944 Seconds  
(without alignments)  
930.966 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MIFPKQYPIINFITAGATVQ.....RFQYIEGEMKTRIRYNRSGA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	955	97.9	267	14	Ricin A chain. Un
2	955	97.9	267	16	Ricin A-chain (RTA
3	955	97.9	290	18	Ricin A-chain ribo
4	955	97.9	290	18	Ricin A-chain ribo
5	955	97.9	332	8	Ricin A. Escheric
6	955	97.9	332	8	Sequence of Ricinu
7	955	97.9	332	10	Ricin A encoded by
8	955	97.9	554	16	Anti-cataract immu
9	955	97.9	562	10	Ricin D. Ricinus

10	955	97.9	565	6	AA250166	Sequence of prepro
11	955	97.9	565	22	AA278300	Castor bean prepro
12	955	97.9	565	22	AA278304	Modified castor be
13	955	97.9	576	8	AA270326	Sequence of Ricinu
14	955	97.9	576	18	AA257877	Castorbean ricin.
15	955	97.9	576	20	AA255892	Castor bean ricin
16	955	97.9	576	21	AA278592	Ricinus communis r
17	955	97.9	576	22	AA278301	Castor bean prepro
18	955	97.9	576	22	AA278302	Castor bean prepro
19	953	97.7	268	14	AA239570	Sequence of ricin-
20	950	97.4	534	14	AA239571	Sequence of G-FIT.
21	950	97.4	565	7	AA260240	Preproicin. Ricci
22	948	97.2	574	8	AA270325	Sequence of Ricinu
23	947	97.1	200	9	AA280164	Biosynthetic multi
24	946	97.0	574	10	AA294793	DNA sequence of ri
25	945	96.9	267	13	AA230722	Ricin A from PIC11
26	945	96.9	267	21	AA219265	Amino acid sequenc
27	944	96.8	332	11	AA206554	Ricin A gene produ
28	938	96.2	267	16	AA274176	Ricin A chain (RTA
29	922	94.6	267	14	AA232430	Ricin A. Syntheti
30	865.5	88.8	540	18	AA25143	Castor oil plant a
31	865.5	88.8	540	18	AA25143	R. communis agglut
32	805.5	82.6	534	8	AA270324	Sequence of Ricin
33	722	74.1	280	10	AA295848	Ricin agglutinin A
34	336	34.5	247	16	AA267359	Trichosanthin anti
35	336	34.5	247	21	AA269048	Amino acid sequenc
36	336	34.5	248	11	AA207518	Synthetic alpha-tr
37	336	34.5	248	13	AA225573	Mature alpha-Trich
38	336	34.5	267	18	AA25140	Trichosanthin (a r
39	336	34.5	267	18	AA221703	Trichosanthin. Tr
40	336	34.5	289	11	AA207514	Trichosanthin from
41	336	34.5	289	13	AA225572	Trichosanthin from
42	336	34.5	289	13	AA229272	Trichosanthin prot
43	336	34.5	289	14	AA232986	Encodes chinese cu
44	336	34.5	289	15	AA25129	Alpha-trichosanthi
45	336	34.5	289	18	AA210468	Chinese cucumber a

ALIGNMENTS

RESULT 1

AA237290  
ID AA237290 standard; protein; 267 AA.

AC AA237290;

XX (updated)

DT 25-MAR-2003 (updated)

DT 09-JAN-2003 (updated)

DT 13-SEP-1993 (first entry)

XX Ricin A chain.

DE Ricin A chain.

XX Type II ribosome-inactivating protein; type II RIP; gelonin;

KW momordin; immunocjugate; autoimmune disease; cell killing; toxin.

XX Unidentified.

OS Unidentified.

XX WO309130-A1.

PN 13-MAY-1993.

PD 04-NOV-1992; 92WO-US09487.

PF 04-NOV-1991; 91US-0787567.

PR 19-JUN-1992; 92US-0901707.

XX (XOMA ) XOMA CORP.

PA Berhard SL, Better MD, Carroll SF, Lane JA, Lei SP;

PI WPI; 1993-167617/20.

XX

PT Analogues of type I ribosome inactivating protein - useful as  
 PT cytotoxic agents, immuno toxins for treating auto immune diseases,  
 PT cancer, graft versus host disease and selective cell killing in-vivo  
 XX  
 PS Claim 1; Page 92; 163pp; English.

XX The invention covers analogues of Type I RIPs. Ricin is a Type II  
 CC RIP whose A chain is homologous to plant type I RIPs. The analogues  
 CC of the invention have a cysteine available for intermolecular  
 CC disulphide bonding at an amino acid position corresp. to a position  
 CC not naturally available for bonding; the cys residue is located in  
 CC the C-terminal region of the analogue between a position corresp. to  
 CC amino acid 251 and the C-terminus of ricin A chain. The analogues are  
 CC pref. joined via a disulphide linkage to a molecule which specifically  
 CC binds to a target cell, e.g. an antibody fragment.  
 CC (Updated on 09-JAN-2003 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 267 AA;

Query Match 97.9%; Score 955; DB 14; Length 267;  
 Best Local Similarity 95.0%; Pred. No. 3.7e-94;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFFKQYPIINFITAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 Db 1 IFFKQYPIINFITAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 60  
 QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQDEDAEAIHTLFTDVQNRVTFAG 111  
 Db 61 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQDEDAEAIHTLFTDVQNRVTFAG 120  
 QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFFIICMISEAR 171  
 Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFFIICMISEAR 180  
 QY 172 FOYIEGEMTRIRYNRRA 190  
 Db 181 FOYIEGEMTRIRYNRRA 199

RESULT 2  
 AAR63902  
 ID AAR63902 standard; protein; 267 AA.

XX AAR63902;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 27-JUL-1995 (first entry)  
 XX Ricin A-chain (RTA).  
 XX Ricin A chain; RTA; ribosome-inactivating proteins; RIPs;  
 KW cytotoxic therapeutic agents; autoimmune disease; cancer;  
 KW graft-versus-host disease.

XX Ricinus communis.  
 XX WO9426910-A1.  
 XX  
 XX 24-NOV-1994.  
 XX  
 XX 12-MAY-1994; 94WO-US05348.  
 XX  
 XX 12-MAY-1993; 93US-0064691.  
 XX (XOMA) XOMA CORP.  
 XX Better MD, Carroll SS, Studnicka GM, Carroll SF;  
 XX WPI; 1995-006804/01.

XX Polynucleotide(s) encoding type I ribosome-inactivating proteins

PT - which are suitable for use as components of cytotoxic  
 PT therapeutic agents.

XX Example 3; Fig 1; 221pp; English.

XX AAR63902 is the ricin A chain gene product, it is analogous to the  
 CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.  
 CC RIPs are the key components of cytotoxic therapeutic agents (CTAs),  
 CC which include gene fusion products and immunoconjugates. CTAs may  
 CC be used to selectively eliminate any cell type to which a RIP  
 CC component is targeted, by the specific binding capacity of the  
 CC second component of the agent. They can be used in the treatment  
 CC of diseases where the elimination of a particular cell type is  
 CC desired, such as autoimmune disease, cancer and graft-versus-host  
 CC disease.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 267 AA;

Query Match 97.9%; Score 955; DB 16; Length 267;  
 Best Local Similarity 95.0%; Pred. No. 3.7e-94;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFFKQYPIINFITAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 Db 1 IFFKQYPIINFITAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 60  
 QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQDEDAEAIHTLFTDVQNRVTFAG 111  
 Db 61 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFPHDPNQDEDAEAIHTLFTDVQNRVTFAG 120  
 QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFFIICMISEAR 171  
 Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFFIICMISEAR 180  
 QY 172 FOYIEGEMTRIRYNRRA 190  
 Db 181 FOYIEGEMTRIRYNRRA 199

RESULT 3  
 AAW25136  
 ID AAW25136 standard; Protein; 290 AA.

XX AAW25136;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 02-DEC-1997 (first entry)  
 XX Ricin A-chain ribosome inhibitory protein inactive precursor.  
 XX Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;  
 KW internal linker; Barley Translation Inhibitor; Trichosanthin;  
 KW Ricin A-chain; Abrin-A A-chain; Saporin; SLT-1; Luffin A; MAP;  
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;  
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;  
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;  
 KW human immunodeficiency virus; acquired immune deficiency syndrome.

XX Synthetic.

XX US5646026-A.

XX 08-JUL-1997.

XX 07-JUN-1995; 95US-0485286.

XX 09-DEC-1992; 92US-0987927.

XX 11-JUN-1990; 90US-0535636.

XX 26-JAN-1995; 95US-0378761.

XX 07-JUN-1995; 95US-0485286.

XX (DOWC) DOWELANCO.



```

XX KW Ricin A; Met-aminopeptidase.
XX OS Escherichia coli.
XX PN EP219237-A.
XX PD 22-APR-1987.
XX PF 19-SEP-1986; 86EP-0307242.
XX PR 06-MAY-1986; 86US-0860330.
XX PR 20-SEP-1985; 85US-0778414.
XX PA (CETU ) CETUS CORP.
XX PI Benbassat A, Bauer KA, Chang S, Chang SY;
XX PI WPI; 1987-110172/16.
XX DR N-PSDB; AAN70152.
XX CC N-terminal methionine free proteins prodn. - by using host
XX CC transformed with vector to express a methionine-amino-peptidase
XX PT Disclosure; Fig. 4; 20pp; English.
XX PS Ricin A may be produced in a form which lacks an N-terminal Met
XX CC using Met-aminopeptidase from E.coli.
XX CC Sequence 332 AA;
XX SQ
Query Match 97.9%; Score 955; DB 8; Length 332;
Best Local Similarity 95.0%; Pred. No. 4.9e-94;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVGRILT-----VLPNRVGLPINQRFILV 51
Db 36 IFPKQYPIINFNTAGATVQSYTNFIRAVGRILTGGADVRHEIPVLPNRVGLPINQRFILV 95
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFAG 111
Db 96 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFAG 155
QY 112 GNYDRLEQLAGNLRENIELGNGLPEALISALYYSTGGTQLPTLARSFFICIQMISEAAR 171
Db 156 GNYDRLEQLAGNLRENIELGNGLPEALISALYYSTGGTQLPTLARSFFICIQMISEAAR 215
QY 172 FOYIEGEMTRIRYNRSA 190
Db 216 FOYIEGEMTRIRYNRSA 234
RESULT 6
ID AAP70838 standard; protein; 332 AA.
XX AAP70838;
XX 25-MAR-2003 (updated)
DT 18-FEB-1991 (first entry)
XX Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)
DE A protein encoded by pR123.
XX Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
XX plant toxin.
XX Ricinus communis.
XX Key Location/Qualifiers
XX Region 1..32
XX FT /note="Leader"
XX FT Region 33..302
XX FT
/notes="A-chain"
/notes="B-chain"
Region
EP237676-A.
23-SEP-1987.
13-NOV-1986; 86EP-0308077.
07-MAR-1986; 86US-0837583.
(CETU ) CETUS CORP.
(CHIR ) CHIRON CORP.
Piatak M;
WPI; 1987-265177/38.
N-PSDB; AAN70519.
New non-glycosylated ricin precursor and toxin etc. - are prepd.
by recombinant DNA procedures with specific isolation steps for
purifier and soluble prods.
Disclosure; Fig 1; 112pp; English.
The full-length sequences encoding ricin A (AAN70520), ricin D
(AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor
then were obtd. using messenger RNA to obtain a cDNA library, and
then probing the library to retrieve the desired cDNA inserts. The
cDNA library was probed using the 35-mer given in AAN70514. Figure 4 (see
AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three
plasmids contg. cDNA inserts obtd. by probing a cDNA library for
sequences encoding ricin B using the probe in AAN70517. The cDNA
inserts can be placed into expression vectors. Site-directed
mutagenesis may be used to place an ATG start codon and a HindIII
site at the beginning of the mature protein (see AAN70518). The
coding sequences of the inserts can be ligated into expression
vectors contg. the phoA promoter-operator and leader sequence
(AAN70523) and suitable retroregulators.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 332 AA;
Query Match 97.9%; Score 955; DB 8; Length 332;
Best Local Similarity 95.0%; Pred. No. 4.9e-94;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVGRILT-----VLPNRVGLPINQRFILV 51
Db 36 IFPKQYPIINFNTAGATVQSYTNFIRAVGRILTGGADVRHEIPVLPNRVGLPINQRFILV 95
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFAG 111
Db 96 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFAG 155
QY 112 GNYDRLEQLAGNLRENIELGNGLPEALISALYYSTGGTQLPTLARSFFICIQMISEAAR 171
Db 156 GNYDRLEQLAGNLRENIELGNGLPEALISALYYSTGGTQLPTLARSFFICIQMISEAAR 215
QY 172 FOYIEGEMTRIRYNRSA 190
Db 216 FOYIEGEMTRIRYNRSA 234
RESULT 7
ID AAP95639 standard; protein; 332 AA.
XX AAP95639;
XX 25-MAR-2003 (updated)
DT 31-OCT-2002 (updated)
XX DT

```



DT 13-AUG-1990 (first entry)  
 XX Ricin A encoded by insert from plasmid pRA123.  
 XX Plasmid pRA123; ricin-A; ricin-B; cytotoxicity.  
 KW Ricinus communis.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH 1..35  
 FT Peptide /label= leader sequence  
 FT Peptide 36..302  
 FT Peptide /label=A-chain  
 FT Peptide 303..314  
 FT Peptide /label=linker  
 FT Peptide 315..332  
 FT Peptide /label=B-chain

XX EP335476-A.

PN 04-OCT-1989.

XX 19-JAN-1989; 89EP-0201162.

XX 08-FEB-1984; 84US-0578115.

XX 08-FEB-1984; 84US-0578121.

XX 09-FEB-1984; 84US-0578122.

XX 07-SEP-1984; 84US-0648759.

XX 20-SEP-1984; 84US-0653515.

XX (CETU ) CETUS CORPORATION.

XX Gelfand D, Lawyer FC, Horn G, Greenfield L, Nitecki D, Kaplan D;  
 PI Piatak MJ;

XX WPI; 1989-286959/40.

XX N-PSDB; AAN91281.

XX Recombinant vectors expressing ricin chains or diphtheria toxin -used for  
 PT prodn. of new immunotoxin conjugates with monoclonal antibodies, having  
 PT high cell specificity and good extracellular stability.

XX Disclosure; Fig 14; 54pp; English.

XX Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for  
 CC ricin A, as well as codons for 12 AAs joining the A to the B chain.

CC Following modification for ease of manipulation the plasmid was used to  
 CC construct expression vectors which express the conjugates in  
 CC host cells.

CC (Updated on 31-OCT-2002 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PF field.)

CC (Updated on 25-MAR-2003 to correct PR field.)

CC (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 332 AA;

Query Match 97.9%; Score 955; DB 10; Length 332;

Best Local Similarity 95.0%; Pred. No. 4.9e-94;

Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINTTAGATVQSYTNFRAVRGLT-----VLPRVGLPINORFLV 51

DB 36 IFPKQYPIINTTAGATVQSYTNFRAVRGLTGDVRRHPIVLPNRVGLPINORFLV 95

QY 52 ELSNHAELSVTLALDVTNAYVGVYAGNSAYFFHPDQEDAEATHLFTDQNYTFAFG 111

DB 96 ELSNHAELSVTLALDVTNAYVGVYAGNSAYFFHPDQEDAEATHLFTDQNYTFAFG 155

QY 112 GNYDRLEQAGLNREINELGNPLEEALISALYYSTGTQTLPTARSFIICQIMISEAR 171

DB 156 GNYDRLEQAGLNREINELGNPLEEALISALYYSTGTQTLPTARSFIICQIMISEAR 215

QY 172 FOYIEGEMTRIRYRRSA 190  
 DB 216 FOYIEGEMTRIRYRRSA 234

XX AAR70827 standard; Protein; 554 AA.

XX AAR70827;

XX 25-MAR-2003 (updated)

XX 31-AUG-1995 (first entry)

XX Anti-cataract immunotoxin.

XX Immunotoxin; heavy chain; light chain; variable region; antibody;

XX ricin-A; cytostatic; cataract; lens opacification; epithelial cell;

XX pHE19; 4197X; monoclonal antibody; MAD.

XX Synthetic.

XX Key Location/Qualifiers

XX 1..27

XX Peptide /label= Sig\_peptide

XX Domain /note= "phoA signal sequence"

XX 28..145

XX Peptide /label= HEAVY

XX 148..166

XX Peptide /note= "Mab 4197X heavy chain"

XX 169..274

XX Domain /label= LINKER

XX 276..544

XX Domain /label= LIGHT

XX 549..554

XX Peptide /note= "Mab 419X light chain"

XX 549..554

XX Peptide /label= RICIN-A

XX 549..554

XX Peptide /label= TAG

XX 549..554

XX Peptide /note= "hexa-histidine tail"

XX WO9503828-A1.

XX 09-FEB-1995.

XX 15-JUL-1994; 94WO-US07919.

XX 02-AUG-1993; 93US-0101329.

XX (HOUS-) HOUSTON BIOTECHNOLOGY INC.

XX Gould RM, Kelleher PJ, Wallace TL, Wood MS;

XX WPI; 1995-082036/11.

XX N-PSDB; AAQ85386.

XX New single chain immuno-toxin - binds specifically to epithelial

XX cells, for inhibiting development of sec. cataracts after

XX extra:capsular cataract extraction.

XX Disclosure; Fig.4; 68pp; English.

XX The immunotoxin given in AAR70827 comprises the heavy and light chain

XX variable regions of anti-lens epithelium IgG3 MAb 4197X linked to

XX ricin-A and a hexa-histidine tag. The DNA construct encoding the

XX immunotoxin was expressed from pHE19 in E. coli.

XX (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 554 AA;

Query Match 97.9%; Score 955; DB 16; Length 554;

Best Local Similarity 95.0%; Pred. No. 9.9e-94;

Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 278 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPEIPVLENRVGLPINQRFILV 337  
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 111  
 DB 338 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 397  
 QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSLFIICQMISEAAR 171  
 DB 398 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSLFIICQMISEAAR 457  
 QY 172 FOYIEGEMRTRIRYNRRSA 190  
 DB 458 FOYIEGEMRTRIRYNRRSA 476  
 RESULT 9  
 AAP90079  
 ID AAP90079 standard; protein; 562 AA.  
 XX  
 AC AAP90079;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Ricin D.  
 XX  
 KW Ricin D; Ricinus communis; castor beans; Zanicariensis variety;  
 KW modified; lectin binding removed; reduced cell binding  
 XX  
 OS Ricinus communis (castor beans).  
 XX  
 PN W08904839-A.  
 XX  
 PD 01-JUN-1989.  
 XX  
 PF 23-NOV-1988; 88WO-US04238.  
 XX  
 PR 24-NOV-1987; 87US-0124735.  
 XX  
 PA (GEMY ) GENETICS INST INC.  
 XX  
 PI Brown EL, Jones S;  
 XX  
 DR WPI; 1989-178366/24.  
 DR N-PSDB; AAN90068.  
 XX  
 PT Modified ricin molecules and toxin conjugates  
 PT - in which the lectin binding function of the B chain  
 PT is removed or diminished to reduce cell binding.  
 XX  
 PS Disclosure; fig 1; 51pp; English.  
 XX  
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment  
 CC of DNA from Ricinus communis, Zanibariensis variety. Patent  
 CC discloses many modifications of ricin in which the lectin binding  
 CC function of the B chain is diminished or removed, and conjugation  
 CC to toxins to eliminate cell binding.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 562 AA;  
 Query Match 97.9%; Score 955; DB 10; Length 562;  
 Best Local Similarity 95.0%; Pred. No. 1e-93;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPEIPVLENRVGLPINQRFILV 95  
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 111

DB 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSLFIICQMISEAAR 171  
 DB 156 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTLARSLFIICQMISEAAR 215  
 QY 172 FOYIEGEMRTRIRYNRRSA 190  
 DB 216 FOYIEGEMRTRIRYNRRSA 234  
 RESULT 10  
 AAP50166  
 ID AAP50166 standard; Protein; 565 AA.  
 XX  
 AC AAP50166;  
 XX  
 DT 16-OCT-1991 (first entry)  
 XX  
 DE Sequence of preprorin encoded by pRCL617.  
 XX  
 KW Toxin; anti-tumour therapy.  
 XX  
 OS Ricinus.  
 XX  
 FH Key  
 FH Peptide Location/Qualifiers  
 FT 1..24  
 FT /label= signal  
 FT Protein 25..565  
 FT Region 292..303  
 FT /label= links the C-terminus of the A chain and  
 FT /label= the N-terminus of the B chain  
 FT Modified-site 34..36  
 FT /label= N-linked glycosylation  
 FT Modified-site 260..262  
 FT /label= N-linked glycosylation  
 FT Modified-site 398..400  
 FT /label= N-linked glycosylation  
 FT Modified-site 438..440  
 FT /label= N-linked glycosylation  
 XX  
 PN EP145111-A.  
 XX  
 PD 19-JUN-1985.  
 XX  
 PF 13-JUL-1984; 84EP-0304801.  
 XX  
 PR 13-MAR-1984; 84GB-0006569.  
 PR 15-JUL-1983; 83GB-0019265.  
 PR 15-JUL-1983; 83CH-0019265.  
 XX  
 PA (UYWA-) UNIV WARWICK.  
 XX  
 PI Lord JM, Roberts LM, Lamb FI;  
 XX  
 DR WPI; 1985-148040/25.  
 DR N-PSDB; AAN50202.  
 XX  
 PT New DNA sequences coding for ricin type plant toxin - or its  
 PT mutants, and modified vectors and host microorganisms  
 XX  
 PS Disclosure; Page 30-30c; 40pp; English.  
 XX  
 CC Preprorin is the whole polypeptide encoded by AAN50202 and the DNA  
 CC encoding this is claimed. Preprorin is obtained from preprorin by  
 CC removal of the AA leader sequence. The linker AA sequence which is  
 CC present in the precursor polypeptide is enzymatically removed in the  
 CC cell to separate the A and B chains, which are joined by a  
 CC disulphide bridge during the formation of the ricin molecule itself.  
 CC This linker region as well as the presumptive amino terminal leader  
 CC or signal sequence are not present in the sequences already  
 CC published by Funatsu et al.

```

XX SQ Sequence 565 AA;
Query Match 97.9%; Score 955; DB 6; Length 565;
Best Local Similarity 95.0%; Pred. No. 1e-93;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHDIPVLPNRVGLPINQRFILV 84
QY 52 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 111
DB 85 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 144
QY 112 GNYDRLEQLAGNLRNLELNGPLEEAI SALTYYSTGTQTLPTLARSFIICIQMISEAAR 171
DB 145 GNYDRLEQLAGNLRNLELNGPLEEAI SALTYYSTGTQTLPTLARSFIICIQMISEAAR 204

QY 172 FOYIEGEMTRIRYNRNSA 190
DB 205 FOYIEGEMTRIRYNRNSA 223

RESULT 11
AAG78300
ID AAG78300 standard; Protein; 565 AA.
XX AC AAG78300;
XX DT 15-NOV-2001 (first entry)
XX DE Castor bean preproricin protein (SEQ ID 1).
XX KW Castor bean plant; preproricin; ricin; A chain; B chain;
XX KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;
XX KW retroviral infection; anti-HIV; virucide activity; viral protease.
XX OS Ricinus communis.
XX FH Location/Qualifiers
XX FT 1..24
XX FT /label= Signal peptide
XX FT Protein
XX FT /label= Ricin_A_chain
XX FT /note= "N-glycosidase"
XX FT 291..302
XX FT /label= Linker peptide
XX FT /note= "Cleaved during activation of ricin"
XX FT 303..565
XX FT /label= Ricin_B_chain
XX FT /note= "Galactose/N-acetylglactosamine-binding lectin"
XX PN WO200160393-A1.
XX PD 23-AUG-2001.
XX PF 15-FEB-2001; 2001WO-US05282.
XX PR 16-FEB-2000; 2000US-0182759.
XX PA (BECH-) BECHTEL BWXT IDAHO LLC.
XX PI Keener WK, Ward TE;
XX DR WPI; 2001-581908/65.
XX DR N-PSDB; AAI64137.
XX PT Novel composition comprising toxin e.g., ricin based antiviral compound
XX PT useful for treating viral infections such as human immunodeficiency
XX PT virus infection.
XX PS Disclosure; Page 47-50; 66pp; English.

```

```

XX CC The sequence relates to preproricin protein encoded by the DNA sequence
XX CC given in AAI64137. The invention relates to a novel toxin (e.g., ricin)
XX CC based antiviral agent which is toxic to virus-infected cells, but
XX CC non-toxic to uninfected cells. The invention has anti-HIV and virucide
XX CC activities. Its mechanism of action is through inactivation of cellular
XX CC ribosomes and enhancement of binding of the antiviral agent to galactose
XX CC residues on cell surfaces, and its cellular internalisation. The
XX CC invention is useful for treating human immunodeficiency virus infection
XX CC and other viral infections, especially retroviral infections. The
XX CC antiviral agent is activated in viral particles or early-stage infected
XX CC cells, killing the cells upon infection and effectively preventing the
XX CC integration of the viral genome into the host genome thereby preventing
XX CC the latency/rebound problem. The agent enters all HIV susceptible cells,
XX CC and not just cells known to act as host cells for the virus. The
XX CC antiviral agent remains inert in a cell until degraded in it, unless the
XX CC cell is infected with the virus, where the viral protease activates it.
XX SQ Sequence 565 AA;

Query Match 97.9%; Score 955; DB 22; Length 565;
Best Local Similarity 95.0%; Pred. No. 1e-93;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHDIPVLPNRVGLPINQRFILV 84
QY 52 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 111
DB 85 ELSNHAELS VTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 144
QY 112 GNYDRLEQLAGNLRNLELNGPLEEAI SALTYYSTGTQTLPTLARSFIICIQMISEAAR 171
DB 145 GNYDRLEQLAGNLRNLELNGPLEEAI SALTYYSTGTQTLPTLARSFIICIQMISEAAR 204
QY 172 FOYIEGEMTRIRYNRNSA 190
DB 205 FOYIEGEMTRIRYNRNSA 223

RESULT 12
AAG78304
ID AAG78304 standard; Protein; 565 AA.
XX AC AAG78304;
XX DT 27-NOV-2001 (first entry)
XX DE Modified castor bean preproricin (SEQ ID 10).
XX DE Castor bean plant; preproricin; ricin; A chain; B chain;
XX KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;
XX KW retroviral infection; anti-HIV; virucide; viral protease.
XX OS Chimeric - Ricinus communis
XX OS Chimeric - Human immunodeficiency virus type 2.
XX FH Location/Qualifiers
XX FT 1..24
XX FT /label= Signal peptide
XX FT Protein
XX FT /label= Proricin
XX FT /note= "Proricin consists of the ricin A chain, a linker
XX FT peptide, and the ricin B chain. Proricin is
XX FT proteolytically cleaved between the A chain and
XX FT the linker to yield mature ricin"
XX FT 25..291
XX FT /label= Ricin_A_chain
XX FT /note= "N-glycosidase"
XX FT 292..303
XX FT /label= Linker_peptide
XX FT Cleavage-site 296..297

```

FT Protein /label= HIV protease\_cleavage\_site  
 FT 304..565  
 FT /label= Ricin B chain  
 FT /note= "Galactose/N-acetylgalactosamine-binding lectin"  
 XX  
 XX WO200160393-A1  
 XX 23-AUG-2001.  
 XX 15-FEB-2001; 2001WO-US05282.  
 XX 16-FEB-2000; 2000US-0182759.  
 XX (BECH-) BECHTEL BWXT IDAHO LLC.  
 XX Keener WK, Ward TE;  
 XX WPI: 2001-581908/65.  
 XX N-PSDB; AA164145.  
 XX Novel composition comprising toxin e.g., ricin based antiviral compound  
 PT useful for treating viral infections such as human immunodeficiency  
 PT virus infection.  
 XX Example 1; Page 59-63; 66pp; English.  
 XX The sequence relates to the amino acid sequence of a modified prepropricin  
 CC protein encoded by AA164145. The invention relates to a novel toxin  
 CC (e.g. ricin) based antiviral agent which is toxic to virus-infected  
 CC cells, but non-toxic to uninfected cells. The invention has anti-HIV and  
 CC virucide activities. The agent is able to enter all HIV susceptible  
 CC cells, and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell unless the cell is infected  
 CC with the HIV virus, where the viral protease activates it. Ricin's  
 CC mechanism of action is through inactivation of cellular ribosomes and  
 CC enhancement of binding of the antiviral agent to galactose residues on  
 CC cell surfaces, and its cellular internalization. The invention is useful  
 CC for treating human immunodeficiency virus infection and other viral  
 CC infections, especially retroviral infections. The antiviral agent is  
 CC activated in viral particles or early-stage infected cells, killing the  
 CC cells upon infection and effectively preventing the integration of the  
 CC viral genome into the host genome thereby preventing the latency/rebound  
 CC problem.  
 XX Sequence 565 AA;  
 XX  
 XX Query Match 97.9%; Score 955; DB 22; Length 565;  
 XX Best Local Similarity 95.0%; Pred. No. 1e-93;  
 XX Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 Db 25 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 84  
 QY 52 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111  
 Db 85 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 144  
 QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEAI SALLYYYSTGGTQLPTLARSFFIICIMISEAAR 171  
 Db 145 GNYDRLEQLAGNLRENIELGNGLPLEEAI SALLYYYSTGGTQLPTLARSFFIICIMISEAAR 204  
 QY 172 FOYIEGEMRTIRYNRNSA 190  
 Db 205 FOYIEGEMRTIRYNRNSA 223  
 RESULT 13  
 AAP70326  
 ID AAP70326 standard; Protein; 576 AA.  
 XX  
 AC AAP70326;  
 XX

DT 25-MAR-2003 (updated)  
 DT 21-MAY-1991 (first entry)  
 XX  
 XX Sequence of Ricinus communis (castor bean) Ricin toxin  
 DE (RT or ricin) E precursor encoded by PRT38.  
 DE  
 XX Lectin; toxin protein; cytotoxic; cytostatic; castor bean;  
 KW plant toxin.  
 XX  
 XX Ricinus communis.  
 XX  
 XX Key Location/Qualifiers  
 FH Region 1..35  
 FT /note= "leader"  
 FT 36..302  
 FT /note= "A-chain"  
 FT 315..576  
 FT /note= "B-chain"  
 XX  
 XX EP237676-A.  
 XX  
 XX 23-SEP-1987.  
 XX  
 XX 13-NOV-1986; 86EP-0308877.  
 XX  
 XX 07-MAR-1986; 86US-0837583.  
 XX  
 XX (CETU ) CETUS CORP.  
 PA (CHIR ) CHIRON CORP.  
 XX  
 XX Piatak M;  
 XX  
 XX WPI: 1987-265177/38.  
 XX N-PSDB; AA70526.  
 XX  
 XX New non-glycosylated ricin precursor and toxin etc. - are prepd.  
 PT by recombinant DNA procedures with specific isolation steps for  
 PT purer and soluble prods.  
 XX  
 XX Disclosure; Fig 14(1-2); 112pp; English.  
 XX  
 XX The full length sequences encoding ricin A (AA70520), ricin D  
 CC (AA70525) putative ricin E (AA70526) and RCA (AA70524) in precursor  
 CC form were obtained, using the messenger RNA to obtain a cDNA library, and  
 CC then probing the library to retrieve the desired cDNA inserts. The  
 CC library was probed using the 35-mer given in AA70514. Figure 4 (see  
 CC AA70520, AA70521, AA70522), shows the nucleotide sequences of three  
 CC plasmids containing cDNA inserts obtained by probing a cDNA library  
 CC for sequences encoding ricin B using the probe in AA70517. The cDNA  
 CC inserts can be placed into expression vectors. Site-directed  
 CC mutagenesis may be used to place an ATG start codon and a HindIII  
 CC site at the beginning of the mature protein, (see AA70518). The  
 CC coding sequences of the inserts can be ligated into expression  
 CC vectors containing the PhOA promoter-operator and leader sequence  
 CC (AA70523) and suitable retroregulators.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX Sequence 576 AA;  
 XX  
 XX Query Match 97.9%; Score 955; DB 8; Length 576;  
 XX Best Local Similarity 95.0%; Pred. No. 1e-93;  
 XX Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 Db 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111  
 Db 96 ELSNHAELS VTLALDVNTNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEAI SALLYYYSTGGTQLPTLARSFFIICIMISEAAR 171

Db 156 GNYDRLEQLAGNLRNENIELGNGPLEEALSAIYYSTGGTQLPTLARSFFIICQMISEAR 215

QY 172 FOYIEGEMTRIRYNRRA 190

Db 216 FOYIEGEMTRIRYNRRA 234

RESULT 14

AAW25787

ID AAW25787 standard; Protein; 576 AA.

XX

AC AAW25787;

DT 25-MAR-2003 (updated)

DT 27-MAR-1998 (first entry)

XX

DE Castor bean ricin.

XX

KW Ricin; cytotoxin; hybrid protein; cell delivery;

KW cell binding ligand; translocation domain; diphtheria toxin B';

KW interleukin-2; T-cell lymphoma; organ rejection; therapy.

XX

OS Ricinus communis.

XX

XX

XX Key Location/Qualifiers

FT Peptide 1..35

FT /label= Sig\_peptide

FT Protein 36..302

FT /label= A-domain

FT Peptide 303..314

FT /label= Linker

FT Domain 315..576

FT /label= B-domain

XX

XX US5668255-A.

XX

XX 16-SEP-1997.

XX

XX 04-AUG-1993; 93US-0102387.

XX

XX 27-JUN-1991; 91US-0722484.

XX 07-JUN-1984; 84US-0618199.

XX 25-APR-1985; 85US-0726808.

XX 07-JUN-1985; 85US-0742554.

XX 22-DEC-1989; 89US-0456095.

XX 14-JUN-1990; 90US-0538276.

XX 04-AUG-1993; 93US-0102387.

XX

XX (SERA-) SERAGEN INC.

XX

XX Murphy JR;

XX

XX WPI; 1997-470103/43.

XX N-PSDB; AAT91638.

XX

XX New hybrid molecules for delivery of agents to cells - comprise a

XX binding domain of a cell binding ligand and a portion of a

XX translocation domain of a protein

XX

XX Example 4; Fig 11A-B; 30pp; English.

XX

XX This polypeptide comprises the castorbean cytotoxin, ricin.

XX DNA (see AAT91638) encoding the enzymatic A domain and a portion

XX of the A-to-B linker peptide of ricin was used to construct a

XX ricin-diphtheria toxin B'-interleukin-2 gene that was expressed in

XX E. coli. The hybrid protein can be isolated and used to treat

XX conditions involving over-production of cells bearing IL2 receptors,

XX such as certain T-cell lymphomas and organ transplant rejection

XX crises. The hybrid inactivates ribosomes in cells bearing IL2

XX receptors, resulting in cessation of protein synthesis and death of

XX target cells. Claimed hybrid proteins comprise a translocation

XX domain and a cell binding domain from e.g. a hormone, growth factor

XX or polypeptide toxin. The hybrid molecules can be used for the

CC delivery of agents (e.g. therapeutic genes, toxins, detectable

CC labels) into cells. The use of a translocation mechanism ensures

CC that the hybrid will be effective in relatively low doses, since a

CC high proportion of the substance of interest will be taken into the

CC targeted cells. The hybrid molecules can be manufactured as a

CC single hybrid recombinant protein, permitting reproducibility,

CC consistency, and the precise control of composition.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 576 AA;

Query Match 97.9%; Score 955; DB 18; Length 576;

Best Local Similarity 95.0%; Pred. No. 1e-93; Indels 10; Gaps 1;

Matches 189; Conservative 0; Mismatches 0;

QY 2 IFPKQYPIINFITAGATVQSNTNFIRAVRGRUT-----VLNVRVGLPINQRFILV 51

Db 36 IFPKQYPIINFITAGATVQSNTNFIRAVRGRUT-----VLNVRVGLPINQRFILV 95

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFPHDPNOEDAEATHLFTDVQNRYYTFAFG 111

Db 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFPHDPNOEDAEATHLFTDVQNRYYTFAFG 155

QY 112 GNYDRLEQLAGNLRNENIELGNGPLEEALSAIYYSTGGTQLPTLARSFFIICQMISEAR 171

Db 156 GNYDRLEQLAGNLRNENIELGNGPLEEALSAIYYSTGGTQLPTLARSFFIICQMISEAR 215

QY 172 FOYIEGEMTRIRYNRRA 190

Db 216 FOYIEGEMTRIRYNRRA 234

RESULT 15

AAW55892

ID AAW55892 standard; Protein; 576 AA.

XX

XX AAW55892;

XX

XX 15-FEB-2000 (first entry)

XX

XX Castor bean ricin toxin.

XX

XX Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria;

XX translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;

XX shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;

XX cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;

XX adipocyte; cancer; virus; infection; antibody.

XX

XX Ricinus communis.

XX

XX US5965406-A.

XX

XX 12-OCT-1999.

XX

XX 07-JUN-1995; 95US-0488246.

XX

XX 04-AUG-1993; 93US-0102387.

XX 07-JUN-1984; 84US-0618199.

XX 27-JUN-1991; 91US-0722484.

XX 25-APR-1985; 85US-0726808.

XX 07-JUN-1985; 85US-0742554.

XX 22-DEC-1989; 89US-0456095.

XX 14-JUN-1990; 90US-0538276.

XX

XX (SERA-) SERAGEN INC.

XX

XX Murphy JR;

XX

XX WPI; 1999-632431/54.

XX N-PSDB; AAZ30663.

XX

XX Recombinant DNA molecule encoding a three part hybrid protein used in

XX the treatment of Aids and genetic deficiency diseases -

XX Example 4; Fig 11; 31pp; English.  
PS The invention relates to a recombinant DNA molecule encoding a hybrid  
XX protein comprising three parts: (a) the first part comprises a portion  
CC of the binding domain of a cell-binding polypeptide ligand allowing the  
CC hybrid protein to bind to an animal cell; (b) the second part comprises  
CC a portion of a translocation domain of a naturally occurring protein  
CC selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera  
CC toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus  
CC toxin, which translocate the third part of the across the cytoplasmic  
CC membrane into the cytosol of the cell; and (c) the third part comprises  
CC a polypeptide entity to be introduced into the cell, which is non-native  
CC to the naturally occurring protein of (b). This sequence represents the  
CC Castor bean ricin toxin sequence for use in generating the hybrid of the  
CC invention. The hybrid molecule enables the direction of appropriate  
CC therapy to affected cells, allowing them to function properly and  
CC alleviate or cure the disease. The hybrid is especially used in treating  
CC genetic deficiency diseases, by delivering to affected cells an enzyme  
CC supplying the missing function, to supplementing cellular levels of a  
CC particular enzyme or a scarce precursor or cofactor, to directing toxins  
CC or other poisons to destroy particular cells (such as adipocytes, cancer  
CC cell, or virus infected-cells), to counteracting viral infections such as  
CC HIV, by introducing appropriate antibodies to viral proteins. It is also  
CC involved in the process of getting non-therapeutic substances such as  
CC detectable labels into cells.  
XX SQ Sequence 576 AA;  
Query Match 97.9%; Score 955; DB 20; Length 576;  
Best Local Similarity 95.0%; Pred. No. 1e-93; Indels 10; Gaps 1;  
Matches 189; Conservative 0; Mismatches 0;  
QY 2 IPPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRYVGLPINQRFILV 51  
Db |||||||  
QY 36 IPPKQYPIINFATTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRYVGLPINQRFILV 95  
Db |||||||  
QY 52 ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDNDQEDAEATHLFTDVQNRYYTFAFG 111  
Db |||||||  
QY 96 ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDNDQEDAEATHLFTDVQNRYYTFAFG 155  
Db |||||||  
QY 112 GNYDRLEQLAGNIRENTELGNGPLEEASALYYVYSTGGTOLPTLARSFIIICMISEAAR 171  
Db |||||||  
QY 156 GNYDRLEQLAGNIRENTELGNGPLEEASALYYVYSTGGTOLPTLARSFIIICMISEAAR 215  
Db |||||||  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db |||||||  
QY 216 FOYIEGEMTRIRYNRRA 234  
Db |||||||

Search completed: February 10, 2004, 16:22:29  
Job time : 32.3944 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 10.9859 Seconds  
(without alignments)  
731.761 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975  
Sequence: 1 MIPKQYPIINFAGATVQ.....RFQYIEGEMTRIRYNR5A 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

1: /cgn2\_6/prodata/1/iaa/5A COMB.pap.\*

2: /cgn2\_6/prodata/1/iaa/5B COMB.pap.\*

3: /cgn2\_6/prodata/1/iaa/6A COMB.pap.\*

4: /cgn2\_6/prodata/1/iaa/6B COMB.pap.\*

5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pap.\*

6: /cgn2\_6/prodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	960	98.5	268	2	US-08-356-786-8
2	960	98.5	534	2	US-08-356-786-10
3	955	97.9	267	1	US-07-901-707-1
4	955	97.9	267	1	US-07-988-430-1
5	955	97.9	267	1	US-08-425-336-1
6	955	97.9	267	1	US-08-488-113B-1
7	955	97.9	267	1	US-08-477-484B-1
8	955	97.9	267	3	US-08-646-360-1
9	955	97.9	267	3	US-08-839-765-1
10	955	97.9	267	3	US-08-136-389-1
11	955	97.9	267	4	US-09-610-838-1
12	955	97.9	267	5	PCT-US92-09487-1
13	955	97.9	290	1	US-08-378-761A-27
14	955	97.9	290	1	US-08-485-286-27
15	955	97.9	290	6	5248606-4
16	945	96.9	267	1	US-08-218-303-16
17	945	96.9	267	2	US-08-338-793D-61
18	945	96.9	267	4	US-09-538-873-1
19	865.5	88.8	540	1	US-08-378-761A-77
20	865.5	88.8	540	1	US-08-485-286-77
21	336	34.5	247	1	US-08-488-113B-6
22	336	34.5	247	1	US-08-477-484B-6
23	336	34.5	247	2	US-08-646-360-6
24	336	34.5	247	3	US-08-839-765-6
25	336	34.5	247	3	US-09-136-389-6
26	336	34.5	247	4	US-09-610-838-6
27	336	34.5	267	1	US-08-378-761A-74

28 336 34.5 267 1 US-08-485-286-74 Sequence 74, Appl  
29 336 34.5 289 1 US-07-923-692C-4 Sequence 4, Appl  
30 336 34.5 289 1 US-08-184-237-4 Sequence 4, Appl  
31 336 34.5 289 2 US-08-482-920-4 Sequence 4, Appl  
32 336 34.5 289 3 US-08-484-341-4 Sequence 4, Appl  
33 336 34.5 289 3 US-08-483-502-4 Sequence 4, Appl  
34 336 34.5 289 4 US-09-726-651A-4 Sequence 4, Appl  
35 320.5 32.9 282 1 US-08-324-301-15 Sequence 15, Appl  
36 310.5 31.8 255 1 US-07-901-707-6 Sequence 6, Appl  
37 310.5 31.8 255 1 US-07-988-430-6 Sequence 6, Appl  
38 310.5 31.8 255 1 US-08-425-336-6 Sequence 6, Appl  
39 310.5 31.8 255 5 PCT-US92-09487-6 Sequence 7, Appl  
40 305 31.3 248 3 US-08-902-486-7 Sequence 2, Appl  
41 305 31.3 290 1 US-08-245-754A-2 Sequence 2, Appl  
42 305 31.3 290 2 US-08-597-731-2 Sequence 15, Appl  
43 305 31.3 496 3 US-08-902-486-15 Sequence 71, Appl  
44 303.5 31.1 250 1 US-08-378-761A-71 Sequence 71, Appl  
45 303.5 31.1 250 1 US-08-485-286-71 Sequence 71, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-356-786-8

; Sequence 8, Application US/08356786

; Patent No. 5877305

; GENERAL INFORMATION:

; APPLICANT: Huston, James S.

; APPLICANT: Oppermann, Hermann

; APPLICANT: Houston, L. L.

; APPLICANT: Ring, David B.

; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer

; TITLE OF INVENTION: Marker

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault

; STREET: Exchange Place, 53 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/356,786

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/831,967

; FILING DATE: 06-FEB-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Pitcher, Edmund R.

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: CRP-053

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 248-7000

; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 268 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-356-786-8

Query Match 98.5%; Score 960; DB 2; Length 268;

Best Local Similarity 95.0%; Pred. No. 2.5e-103;

Matches 190; Conservative 0; Mismatches 0; Indels 10; Gaps 1;



QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50  
 |||||  
 Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFIL 60  
 |||||  
 QY 51 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 110  
 |||||  
 Db 61 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 120  
 |||||  
 QY 111 GGNVDRLEQLAGNLRNIELGNPLSEASALYYSTGGTQPLTLARSFFIICQMISEAA 170  
 |||||  
 Db 121 GGNVDRLEQLAGNLRNIELGNPLSEASALYYSTGGTQPLTLARSFFIICQMISEAA 180  
 |||||  
 QY 171 RFOYIEGEMRTRIRYNRRA 190  
 |||||  
 Db 181 RFOYIEGEMRTRIRYNRRA 200  
 |||||

## RESULT 2

US-08-356-786-10  
 ; Sequence 10, Application US/08356786  
 ; Patent No. 587305  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Huston, James S.  
 ; APPLICANT: Oppermann, Hermann  
 ; APPLICANT: Houston, L. L.  
 ; APPLICANT: King, David B.  
 ; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
 ; STREET: Exchange Place, 53 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/356,786  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/831,967  
 FILING DATE: 06-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pitcher, Edmund R.  
 REGISTRATION NUMBER: 27,829  
 REFERENCE/DOCKET NUMBER: CRP-053  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 248-7000  
 TELEFAX: (617) 248-7100  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 534 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-356-786-10

Query Match 98.5%; Score 960; DB 2; Length 534;  
 Best Local Similarity 95.0%; Pred. No. 6.8e-103;  
 Matches 190; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50  
 |||||  
 Db 3 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFIL 62  
 |||||  
 QY 51 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 110  
 |||||

Db 63 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 122  
 |||||  
 QY 111 GGNVDRLEQLAGNLRNIELGNPLSEASALYYSTGGTQPLTLARSFFIICQMISEAA 170  
 |||||  
 Db 123 GGNVDRLEQLAGNLRNIELGNPLSEASALYYSTGGTQPLTLARSFFIICQMISEAA 182  
 |||||  
 QY 171 RFOYIEGEMRTRIRYNRRA 190  
 |||||  
 Db 183 RFOYIEGEMRTRIRYNRRA 202  
 |||||

## RESULT 3

US-07-901-707-1  
 ; Sequence 1, Application US/07901707  
 ; Patent No. 5376546  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bernhardt, Susan L.  
 ; APPLICANT: Better, Marc D.  
 ; APPLICANT: Carroll, Steve F.  
 ; APPLICANT: Lane, Julie A.  
 ; TITLE OF INVENTION: Materials Comprising and Methods of  
 ; NUMBER OF SEQUENCES: 57  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
 ; ADDRESSEE: Bicknell  
 ; STREET: Two First National Plaza, 20 South Clark  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60603

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/901,707  
 FILING DATE: 19920619  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/787,567  
 FILING DATE: 04-NOV-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 5376546and, Greta E.  
 REGISTRATION NUMBER: 35,302  
 REFERENCE/DOCKET NUMBER: 27129/30910  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 346-5750  
 TELEFAX: (312) 984-5750  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 267 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-901-707-1

Query Match 97.9%; Score 955; DB 1; Length 267;  
 Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 51  
 |||||  
 Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFIL 60  
 |||||  
 QY 52 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 111  
 |||||  
 Db 61 VELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDNQDEAEATHLFTDVQNRYYTFAF 120  
 |||||  
 QY 112 GGNVDRLEQLAGNLRNIELGNPLSEASALYYSTGGTQPLTLARSFFIICQMISEAA 171  
 |||||

Db 121 GNYDRLEQAGNLRENIELGNGLPEEAISALYYSTGGTQLTPLARSFICQMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199

RESULT 4  
US-07-988-430-1  
; Sequence 1, Application US/07988430  
; Patent No. 5416202  
; GENERAL INFORMATION:  
; APPLICANT: Bernhard, Susan L.  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Lane, Julie A.  
; APPLICANT: Lei, Shau-Ping  
; TITLE OF INVENTION: Materials Comprising and Methods of  
; Preparation and Use for Ribosome-Inactivating Proteins  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/988,430  
; FILING DATE: 19921209  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5416202and, Greta E.  
; REGISTRATION NUMBER: 35302  
; REFERENCE/DOCKET NUMBER: 31133  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-988-430-1

Query Match 97.9%; Score 955; DB 1; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120

QY 112 GNYDRLEQAGNLRENIELGNGLPEEAISALYYSTGGTQLTPLARSFICQMISEAAR 171  
Db 121 GNYDRLEQAGNLRENIELGNGLPEEAISALYYSTGGTQLTPLARSFICQMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199

RESULT 5  
US-08-425-336-1  
; Sequence 1, Application US/08425336  
; Patent No. 5621083  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; Proteins  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/425,336  
; FILING DATE: 18-APR-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/064,691  
; FILING DATE: 12-MAY-1993  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Thomas C.  
; REGISTRATION NUMBER: P-36,989  
; REFERENCE/DOCKET NUMBER: 31394  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-425-336-1

Query Match 97.9%; Score 955; DB 1; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQAGNLRENIELGNGLPEEAISALYYSTGGTQLTPLARSFICQMISEAAR 171

Db 121 GNYDRLEQAGNLRNIELGNGLPEEAISALYYSTGGTQLPTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199  
RESULT 6  
US-08-488-113B-1  
; Sequence 1, Application US/08488113B  
; Patent No. 5744580  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,113B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-488-113B-1  
Query Match 97.9%; Score 955; DB 1; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQVPLINFTTAGTQVSYTNFIRAVRGELT-----VLNRYVGLPINQRFILV 51  
Db 1 IFPKQVPLINFTTAGTQVSYTNFIRAVRGELT-----VLNRYVGLPINQRFILV 60

QY 52 ELSNHAELSVTIALDVTNAYVYGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELSVTIALDVTNAYVYGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQAGNLRNIELGNGLPEEAISALYYSTGGTQLPTLARSFIIQIMISEAAR 171  
Db 121 GNYDRLEQAGNLRNIELGNGLPEEAISALYYSTGGTQLPTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRSA 190  
Db 181 FOYIEGEMTRIRYNRSA 199  
RESULT 7  
US-08-477-484B-1  
; Sequence 1, Application US/08477484B  
; Patent No. 5756699  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,484B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-477-484B-1

Query Match

97.9%; Score 955; DB 1; Length 267;

Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
Qy 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHPIVLPNRVGLPINQRFILV 60  
Qy 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 120  
Qy 112 GNYDRLEQLAGNLRNIELGNGLPDEEALSALYYSTGGTQTLPTLARSFIICMISEAAAR 171  
Db 121 GNYDRLEQLAGNLRNIELGNGLPDEEALSALYYSTGGTQTLPTLARSFIICMISEAAAR 180  
Qy 172 FOYIEGEMRTRIRYNRRA 190  
Db 181 FOYIEGEMRTRIRYNRRA 199  
RESULT 8  
US-08-646-360-1  
; Sequence 1, Application US/08646360  
; Patent No. 5837491  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/646,360  
; FILING DATE: 13-MAY-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05348  
; FILING DATE: 12-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 200-70.P4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-646-360-1  
Query Match 97.9%; Score 955; DB 2; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
Qy 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHPIVLPNRVGLPINQRFILV 60  
Qy 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFG 120  
Qy 112 GNYDRLEQLAGNLRNIELGNGLPDEEALSALYYSTGGTQTLPTLARSFIICMISEAAAR 171  
Db 121 GNYDRLEQLAGNLRNIELGNGLPDEEALSALYYSTGGTQTLPTLARSFIICMISEAAAR 180  
Qy 172 FOYIEGEMRTRIRYNRRA 190  
Db 181 FOYIEGEMRTRIRYNRRA 199  
RESULT 9  
US-08-839-765-1  
; Sequence 1, Application US/08839765  
; Patent No. 6145631  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/839,765  
; FILING DATE: 15-APR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-765-1

Query Match 97.9%; Score 955; DB 3; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLTTGADVVRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEAI S A L Y Y S T G G T Q L P T L A R S F I C I Q M I S E A A R 171  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEAI S A L Y Y S T G G T Q L P T L A R S F I C I Q M I S E A A R 180  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 181 FOYIEGEMTRIRYNRRA 199

## RESULT 10

US-09-136-389-1  
Sequence 1, Application US/09136389  
Patent No. 6146850  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-136-389-1

Query Match 97.9%; Score 955; DB 3; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLTTGADVVRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEAI S A L Y Y S T G G T Q L P T L A R S F I C I Q M I S E A A R 171  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEAI S A L Y Y S T G G T Q L P T L A R S F I C I Q M I S E A A R 180  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 181 FOYIEGEMTRIRYNRRA 199

## RESULT 11

US-09-610-838-1  
Sequence 1, Application US/09610838  
Patent No. 6376217  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/610,838  
FILING DATE: 06-JUL-2000  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE: 18-AUG-1998  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348

; FILING DATE: 12-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 200-70.P4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-610-838-1

Query Match 97.9%; Score 955; DB 4; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
Qy 2 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 60  
  
Qy 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 120  
  
Qy 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYVSTGTGTLPTLARSFIICQMISEAAR 171  
Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYVSTGTGTLPTLARSFIICQMISEAAR 180  
  
Qy 172 FOYIEGEMTRIRYNRRA 190  
Db 181 FOYIEGEMTRIRYNRRA 199

RESULT 12  
PCT-US92-09487-1  
; Sequence 1, Application PC/TUS9209487  
; GENERAL INFORMATION:  
; APPLICANT: Bernhard, Susan L.  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Lane, Julie A.  
; APPLICANT: Lei, Shau-Ping  
; TITLE OF INVENTION: Materials Comprising and Methods of Preparation and Use for Ribosome-Inactivating Proteins  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/09487  
; FILING DATE: 19921104  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Noland, Greta E.  
; REGISTRATION NUMBER: 35302  
; REFERENCE/DOCKET NUMBER: 31133  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US92-09487-1

Query Match 97.9%; Score 955; DB 5; Length 267;  
Best Local Similarity 95.0%; Pred. No. 9.4e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
Qy 2 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFETAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 60  
  
Qy 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEALTHLFTDVQNYTFAFG 120  
  
Qy 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYVSTGTGTLPTLARSFIICQMISEAAR 171  
Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYVSTGTGTLPTLARSFIICQMISEAAR 180  
  
Qy 172 FOYIEGEMTRIRYNRRA 190  
Db 181 FOYIEGEMTRIRYNRRA 199

RESULT 13  
US-08-378-761A-27  
; Sequence 27, Application US/08378761A  
; Patent No. 5635384  
; GENERAL INFORMATION:  
; APPLICANT: WALSH, TERENCE A  
; APPLICANT: HEY, TIMOTHY D  
; APPLICANT: MORGAN, ALICE ER  
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ANDREA T. BORUCKI  
; STREET: 9330 ZIONSVILLE ROAD  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: US  
; ZIP: 46268  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-27

Query Match 97.9%; Score 955; DB 1; Length 290;  
Best Local Similarity 95.0%; Pred. No. 1.1e-102;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
QY 2 IPFKQYPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 25 IPFKQYPIINFITAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
  
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 111  
Db 85 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 144  
  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFICIMISEAAR 171  
Db 145 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFICIMISEAAR 204  
  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 205 FOYIEGEMTRIRYNRRA 223

RESULT 14  
US-08-485-286-27  
Sequence 27, Application US/08485286  
Patent No. 5646026  
Patent No. 5646026 5646119  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
TITLE OF INVENTION: USING  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,286  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/378761  
FILING DATE: 26-JAN-1995

ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-286-27

Query Match 97.9%; Score 955; DB 1; Length 290;  
Best Local Similarity 95.0%; Pred. No. 1.1e-102;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
QY 2 IPFKQYPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 25 IPFKQYPIINFITAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
  
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 111  
Db 85 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 144  
  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFICIMISEAAR 171  
Db 145 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFICIMISEAAR 204  
  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 205 FOYIEGEMTRIRYNRRA 223

RESULT 15  
5248606-4  
Patent No. 5248606  
APPLICANT: WALSH, TERENCE A.; HEY, TIMOTHY D.; MORGAN,  
ALICE E.R.  
TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND  
ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATIN  
NUMBER OF SEQUENCES: 49  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/535,636  
FILING DATE: 11-JUN-1990  
SEQ ID NO: 4:  
LENGTH: 290  
5248606-4

Query Match 97.9%; Score 955; DB 6; Length 290;  
Best Local Similarity 95.0%; Pred. No. 1.1e-102;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
  
QY 2 IPFKQYPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 25 IPFKQYPIINFITAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
  
QY 52 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 111  
Db 85 ELSNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAI THLFTDVQNRVTFAG 144  
  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFICIMISEAAR 171  
Db 145 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFICIMISEAAR 204  
  
QY 172 FOYIEGEMTRIRYNRRA 190  
Db 205 FOYIEGEMTRIRYNRRA 223

Search completed: February 10, 2004, 16:29:35

Sun Feb 15 07:29:57 2004

us-10-083-336a-11.rai

Page 9

Job time : 10.9859 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:26:46 ; Search time 24.507 Seconds  
(without alignments)  
1623.314 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MIFPKQYPIINFITAGATVQ.....RFQYIEGEMRTRINRNSA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	975	100.0	190	12	US-10-083-336A-11
2	971	99.6	189	12	US-10-083-336A-6
3	966	99.1	188	12	US-10-083-336A-4
4	960	98.5	200	12	US-10-083-336A-10
5	956	98.1	188	12	US-10-083-336A-8
6	956	98.1	199	12	US-10-083-336A-5
7	955	97.9	267	12	US-10-127-890-1
8	955	97.9	576	12	US-10-083-336A-1
9	951	97.5	198	12	US-10-083-336A-3
10	945	96.9	267	12	US-10-282-935-1
11	945	96.9	267	12	US-10-440-796-1
12	941	96.5	198	12	US-10-083-336A-7
13	939.5	96.4	185	12	US-10-083-336A-9
14	679	69.6	179	12	US-10-083-336A-2
15	336	34.5	247	10	US-09-792-793A-39

16	336	34.5	247	12	US-10-127-890-6	Sequence 6, Appli
17	336	34.5	247	12	US-10-375-209A-39	Sequence 39, Appli
18	336	34.5	289	12	US-10-280-679B-4	Sequence 4, Appli
19	305	31.3	247	10	US-09-792-793A-34	Sequence 34, Appli
20	305	31.3	247	12	US-10-375-209A-34	Sequence 34, Appli
21	297.5	30.5	251	12	US-10-282-935-3	Sequence 3, Appli
22	297.5	30.5	251	12	US-10-440-796-3	Sequence 3, Appli
23	273	28.0	263	12	US-10-127-890-7	Sequence 7, Appli
24	272	27.9	263	12	US-10-127-890-4	Sequence 4, Appli
25	257.5	26.4	248	12	US-10-127-890-5	Sequence 5, Appli
26	254	26.1	252	9	US-09-347-064-2	Sequence 2, Appli
27	254	26.1	252	9	US-09-347-064-8	Sequence 8, Appli
28	248.5	25.5	251	12	US-10-127-890-107	Sequence 107, App
29	247.5	25.4	251	12	US-10-127-890-106	Sequence 106, App
30	247.5	25.4	251	12	US-10-127-890-110	Sequence 110, App
31	247.5	25.4	251	12	US-10-127-890-111	Sequence 111, App
32	246.5	25.3	251	9	US-09-765-527-247	Sequence 247, App
33	246.5	25.3	251	12	US-10-127-890-2	Sequence 2, Appli
34	246.5	25.3	251	12	US-10-127-890-99	Sequence 99, Appli
35	246.5	25.3	251	12	US-10-127-890-100	Sequence 100, App
36	246.5	25.3	251	12	US-10-127-890-101	Sequence 101, App
37	246.5	25.3	251	12	US-10-127-890-102	Sequence 102, App
38	246.5	25.3	251	12	US-10-127-890-103	Sequence 103, App
39	246.5	25.3	251	12	US-10-127-890-104	Sequence 104, App
40	246.5	25.3	251	12	US-10-127-890-105	Sequence 105, App
41	246.5	25.3	316	12	US-10-074-596-1	Sequence 1, Appli
42	246.5	25.3	507	12	US-10-074-596-11	Sequence 11, Appli
43	245.5	25.2	293	9	US-09-765-527-259	Sequence 259, App
44	245.5	25.2	309	9	US-09-765-527-253	Sequence 253, App
45	245.5	25.2	332	9	US-09-765-527-251	Sequence 251, App

#### ALIGNMENTS

RESULT 1  
US-10-083-336A-11  
; Sequence 11, Application US/10083336A  
; Publication No. US20030181665A1  
; GENERAL INFORMATION:  
; APPLICANT: Olson, Mark A  
; APPLICANT: Millard, Charles B  
; APPLICANT: Byrne, Michael P  
; APPLICANT: Wannemacher, Robert W  
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
; FILE REFERENCE: P67452US0 (RIID 01-58)  
; CURRENT APPLICATION NUMBER: US/10/083,336A  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; TYPE: PRT  
; ORGANISM: Ricinus communis  
US-10-083-336A-11

Query Match	100.0%;	Score	975;	DB	12;	Length	190;
Best Local Similarity	100.0%;	Pred. No.	1.1e-103;				
Matches	190;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MIFPKQYPIINFITAGATVQSYTNFRAVGRITVLPNRVGLPINQRFILVELSNHAELS	60				
Db	1	MIFPKQYPIINFITAGATVQSYTNFRAVGRITVLPNRVGLPINQRFILVELSNHAELS	60				
Qy	61	VTALDVNTNAYVGVYRAGNSAYFFHPDNQSDAEATHLFTDVQNRYYTFAFGNYYDRLEQL	120				
Db	61	VTALDVNTNAYVGVYRAGNSAYFFHPDNQSDAEATHLFTDVQNRYYTFAFGNYYDRLEQL	120				
Qy	121	AGNLRENIELNGPLEEAISALYYSTGCTQLTLARSFFICIMISEAARFQYIEGEMR	180				
Db	121	AGNLRENIELNGPLEEAISALYYSTGCTQLTLARSFFICIMISEAARFQYIEGEMR	180				
Qy	181	TRIRYNRNSA	190				

```
Db 181 TRIRYNRS 190
|||||
1 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELSV 60
62 TLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTFAGGNYDRLEQLA 121
|||||
61 TLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTFAGGNYDRLEQLA 120
|||||
122 GNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEARFOYIEGEMRT 181
|||||
121 GNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEARFOYIEGEMRT 180
|||||
182 RIRYNRS 189
|||||
181 RIRYNRS 188
|||||
RESULT 4
US-10-083-336A-10
; Sequence 10, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-6
Query Match 99.6%; Score 971; DB 12; Length 189;
Best Local Similarity 100.0%; Pred. No. 3.1e-103; Mismatches 0; Indels 0; Gaps 0;
Matches 189; Conservative 0;
Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELS 60
|||||
1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELS 60
|||||
61 VTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTFAGGNYDRLEQL 120
|||||
61 VTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTFAGGNYDRLEQL 120
|||||
121 AGNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEARFOYIEGEMR 180
|||||
121 AGNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEARFOYIEGEMR 180
|||||
181 TRIRYNRS 189
|||||
181 TRIRYNRS 189
|||||
RESULT 3
US-10-083-336A-4
; Sequence 4, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-4
Query Match 99.1%; Score 966; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.1e-102; Mismatches 0; Indels 0; Gaps 0;
Matches 188; Conservative 0;
Db 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELSV 61
|||||
1 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFILVELSNHAELSV 60
|||||
62 TLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTFAGGNYDRLEQLA 121
|||||
61 TLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTFAGGNYDRLEQLA 120
|||||
122 GNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEARFOYIEGEMRT 181
|||||
121 GNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEARFOYIEGEMRT 180
|||||
182 RIRYNRS 189
|||||
181 RIRYNRS 188
|||||
RESULT 5
US-10-083-336A-8
; Sequence 8, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; QUERY MATCH 98.5%; SCORE 960; DB 12; LENGTH 200;
; BEST LOCAL SIMILARITY 95.0%; PRED. NO. 6.1e-102; MISMATCHES 0; INDELS 10; GAPS 1;
; MATCHES 190; CONSERVATIVE 0;
Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFIL 50
|||||
1 MIFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTVLPNRVGLPINQRFIL 60
|||||
51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTF 110
|||||
61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEAIHLFTDVQNRVTF 120
|||||
111 GGNVDRLEQLAGNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEAA 170
|||||
121 GGNVDRLEQLAGNLRENIELGNGLPBEAISALYYSTGGTQLPTLARSFIIQMISEAA 180
|||||
171 RFQYIEGEMRTIRYNRS 190
|||||
181 RFQYIEGEMRTIRYNRS 200
|||||
US-10-083-336A-8
; Sequence 8, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
```

```

; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
; US-10-083-336A-8

Query Match      98.1%; Score 956; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e-101;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PKQYPIINFNTAGATVQSYTNFIRAVGRGLTVPNVRGLPINQRFILVELSNAHAELSVTL 63
Db 3 PKQYPIINFNTAGATVQSYTNFIRAVGRGLTVPNVRGLPINQRFILVELSNAHAELSVTL 62

QY 64 ALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVQRYTFAGGNYDLEQLAGN 123
Db 63 ALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVQRYTFAGGNYDLEQLAGN 122

QY 124 LRENIELGNGLPLEEASALYYSTGGTQLPFLARSFICIQMISEAARFOYIEGEMTRI 183
Db 123 LRENIELGNGLPLEEASALYYSTGGTQLPFLARSFICIQMISEAARFOYIEGEMTRI 182

QY 184 RYNRRS 189
Db 183 RYNRRS 188

RESULT 6
US-10-083-336A-5
; Sequence 5, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083, 336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Ricinus communis
; US-10-083-336A-5

Query Match      98.1%; Score 956; DB 12; Length 199;
Best Local Similarity 95.0%; Pred. No. 1.7e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFNTAGATVQSYTNFIRAVGRGLT-----VLPNVRGLPINQRFIL 50
Db 1 MIFPKQYPIINFNTAGATVQSYTNFIRAVGRGLTGDVVRHEIPVLPNVRGLPINQRFIL 60

QY 51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVQRYTF 110
Db 61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVQRYTF 120

QY 111 GGNDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLPFLARSFICIQMISEA 170
Db 121 GGNDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLPFLARSFICIQMISEA 180

QY 171 RFQYIEGEMTRIYNRRS 189
Db 181 RFQYIEGEMTRIYNRRS 199

RESULT 7
US-10-127-890-1
; Sequence 1, Application US/10127890
; Publication No. US20030166196A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Better, Marc D.
; Carroll, Stephen F.
; Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-10-127-890-1

Query Match      97.9%; Score 955; DB 12; Length 267;
Best Local Similarity 95.0%; Pred. No. 3.5e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVGRGLT-----VLPNVRGLPINQRFILV 51
Db 1 IFPKQYPIINFNTAGATVQSYTNFIRAVGRGLTGDVVRHEIPVLPNVRGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVQRYTF 111
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDQVQRYTF 120

QY 112 GGNDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLPFLARSFICIQMISEA 171
Db 121 GGNDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLPFLARSFICIQMISEA 180

QY 172 RFQYIEGEMTRIYNRRS 190
Db 181 RFQYIEGEMTRIYNRRS 199

```

```
RESULT 8
US-10-083-336A-1
; Sequence 1, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-1

Query Match          97.9%; Score 955; DB 12; Length 576;
Best Local Similarity 95.0%; Pred. No. 1e-100;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQVPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 36 IFPKQVPIINFNTAGATVQSYTNFIRAVRGRLTGGADVREHIEPVLPNRVGLPINQRFILV 95
QY 52 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOQDEAEATHLFTDVQNRYYTFAFG 111
DB 96 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOQDEAEATHLFTDVQNRYYTFAFG 155
QY 112 GNYDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQLTPLARSFFIICQMISEAAR 171
DB 156 GNYDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQLTPLARSFFIICQMISEAAR 215
QY 172 FOYIEGEMTRIRYNNRSA 190
DB 216 FOYIEGEMTRIRYNNRSA 234

RESULT 9
US-10-083-336A-3
; Sequence 3, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-3

Query Match          97.5%; Score 951; DB 12; Length 198;
Best Local Similarity 94.9%; Pred. No. 6.5e-101;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQVPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 1 IFPKQVPIINFNTAGATVQSYTNFIRAVRGRLTGGADVREHIEPVLPNRVGLPINQRFILV 60
QY 52 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOQDEAEATHLFTDVQNRYYTFAFG 111
DB 111 ELSNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOQDEAEATHLFTDVQNRYYTFAFG 182

RESULT 10
US-10-282-935-1
; Sequence 1, Application US/10282935
; Publication No. US20030143193A1
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETTIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN F.
; APPLICANT: BALUNA, ROKANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF
; FILE REFERENCE: UTSD:884US
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 09/538,873
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/126,826
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-282-935-1

Query Match          96.9%; Score 945; DB 12; Length 267;
Best Local Similarity 94.9%; Pred. No. 4.9e-100;
Matches 187; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 4 PKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 53
DB 3 PKQYPIINFNTAGATVQSYTNFIRAVRGRLTGGADVREHIEPVLPNRVGLPINQRFILV 62
QY 54 SNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOQDEAEATHLFTDVQNRYYTFAFG 113
DB 63 SNHAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOQDEAEATHLFTDVQNRYYTFAFG 122
QY 114 YDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQLTPLARSFFIICQMISEAARFQ 173
DB 123 YDRLEQLAGNLENIELGNGLPLEEALISALYYSTGGTQLTPLARSFFIICQMISEAARFQ 182
QY 174 YIEGEMTRIRYNNRSA 190
DB 183 YIEGEMTRIRYNNRSA 199

RESULT 11
US-10-440-796-1
; Sequence 1, Application US/10440796
; Publication No. US20040009148A1
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETTIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN F.
; APPLICANT: BALUNA, ROKANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; FILE REFERENCE: UTSD:603
```

CURRENT APPLICATION NUMBER: US/10/440,796  
CURRENT FILING DATE: 2003-05-19  
PRIOR APPLICATION NUMBER: US/09/538,873  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: 60/126,826  
PRIOR FILING DATE: 1999-03-30  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide  
US-10-440-796-1

Query Match 96.9%; Score 945; DB 12; Length 267;  
Best Local Similarity 94.9%; Pred. No. 4.9e-100;  
Matches 187; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 4 PKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILVEL 53  
DB 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFILVEL 62  
QY 54 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 113  
DB 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 122  
QY 114 YDRLEQLAGNRENTLGNPLEEALISALYYSTGTQTLPTLARSFIICQMISEAARFQ 173  
DB 123 YDRLEQLAGNRENTLGNPLEEALISALYYSTGTQTLPTLARSFIICQMISEAARFQ 182  
QY 174 YIEGEMTRIRYNRRA 190  
DB 183 YIEGEMTRIRYNRRA 199

RESULT 12  
US-10-083-336A-7  
Sequence 7, Application US/10083336A  
Publication No. US20030181665A1  
GENERAL INFORMATION:  
APPLICANT: Olson, Mark A  
APPLICANT: Millard, Charles B  
APPLICANT: Byrne, Michael P  
APPLICANT: Wannemacher, Robert W  
TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
FILE REFERENCE: P67452US0 (RIID 01-58)  
CURRENT APPLICATION NUMBER: US/10/083,336A  
CURRENT FILING DATE: 2002-05-21  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 198  
TYPE: PRT  
ORGANISM: Ricinus communis  
US-10-083-336A-7  
Query Match 96.5%; Score 941; DB 12; Length 198;  
Best Local Similarity 94.9%; Pred. No. 9.2e-100;  
Matches 186; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 4 PKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILVEL 53  
DB 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFILVEL 62  
QY 54 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 113  
DB 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGN 122  
QY 114 YDRLEQLAGNRENTLGNPLEEALISALYYSTGTQTLPTLARSFIICQMISEAARFQ 173

DB 123 YDRLEQLAGNRENTLGNPLEEALISALYYSTGTQTLPTLARSFIICQMISEAARFQ 182  
QY 174 YIEGEMTRIRYNRRA 189  
DB 183 YIEGEMTRIRYNRRA 198  
RESULT 13  
US-10-083-336A-9  
Sequence 9, Application US/10083336A  
Publication No. US20030181665A1  
GENERAL INFORMATION:  
APPLICANT: Olson, Mark A  
APPLICANT: Millard, Charles B  
APPLICANT: Byrne, Michael P  
APPLICANT: Wannemacher, Robert W  
TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
FILE REFERENCE: P67452US0 (RIID 01-58)  
CURRENT APPLICATION NUMBER: US/10/083,336A  
CURRENT FILING DATE: 2002-05-21  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 185  
TYPE: PRT  
ORGANISM: Ricinus communis  
US-10-083-336A-9  
Query Match 96.4%; Score 939.5; DB 12; Length 185;  
Best Local Similarity 98.4%; Pred. No. 1.2e-99;  
Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTVLPNRVGLPINQRFILVELSNHAELSV 61  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT---NRVGLPINQRFILVELSNHAELSV 57  
QY 62 TLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGNDRLEOLA 121  
DB 58 TLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAGGNDRLEOLA 117  
QY 122 GNRENTLGNPLEEALISALYYSTGTQTLPTLARSFIICQMISEAARFQYIEGEMRT 181  
DB 118 GNRENTLGNPLEEALISALYYSTGTQTLPTLARSFIICQMISEAARFQYIEGEMRT 177  
QY 182 RIRYNRRA 189  
DB 178 RIRYNRRA 185

RESULT 14  
US-10-083-336A-2  
Sequence 2, Application US/10083336A  
Publication No. US20030181665A1  
GENERAL INFORMATION:  
APPLICANT: Olson, Mark A  
APPLICANT: Millard, Charles B  
APPLICANT: Byrne, Michael P  
APPLICANT: Wannemacher, Robert W  
TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
FILE REFERENCE: P67452US0 (RIID 01-58)  
CURRENT APPLICATION NUMBER: US/10/083,336A  
CURRENT FILING DATE: 2002-05-21  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 179  
TYPE: PRT  
ORGANISM: Ricinus communis  
US-10-083-336A-2  
Query Match 69.6%; Score 679; DB 12; Length 179;  
Best Local Similarity 93.1%; Pred. No. 9.9e-70;  
Matches 134; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
US-10-083-336A-2

```
QY      2  IFPKQYPLINFTTAGATVQSYTNEIFRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db      36  IFPKQYPLINFTTAGATVQSYTNEIFRAVRGRLTGDVRHEIPVLEPNRVGLPINQRFILV 95
QY     52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQNOEDAETHLFTDQVQRYTFAG 111
Db     96  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQNOEDAETHLFTDQVQRYTFAG 155
QY    112  GNYDRLEQLAGNLRNIELGNGPL 135
Db    156  GNYDRLEQLAGNLRNIELGNGPL 179
```

## RESULT 15

```
US-09-792-793A-39
; Sequence 39, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggins, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Trichosantheus kirilowii
US-09-792-793A-39
```

```
Query Match      34.5%; Score 336; DB 10; Length 247;
Best Local Similarity 38.9%; Pred. No. 3.9e-30;
Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY      10  INFETAGATVQSYTNEIFRAVRGRLTVLPN-----RVGLPINQRFILVELSNHAEL 59
Db      2  VSRLSGATSSSYGVFISNR---KALPNERKLYDIPLLRSSLPQSQRVALIHLTNVADE 58
QY     60  SVTLALDVTNAYVVGVRAGNSAYFFHPDQNOEDA-EAITHLFTDQVQRYTFAGGNYDRLE 118
Db     59  TISVAIDVTNVMGYVRAGDTSYFF--NEASATEAAKVVKDAMRKVTLPYSGNYERLQ 115
QY    119  QLAGNLRNIELGNGPLEEASALYYVYSTGQTLPILARSFIICQMISEAARFQYIEGE 178
Db    116  TAAGKIRENIEPLGLPALDSAITLTFYNN-----SAASALMVLIOSTSEARFYKFEQQ 170
QY    179  MRTRI 183
Db    171  IGRV 175
```

Search completed: February 10, 2004, 16:53:56  
Job time : 25.507 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:35 ; Search time 10.2817 Seconds  
(without alignments)  
1777.145 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MIFPKQYPIINFITAGATVQ.....RFQYIEGEMRTIRYNRSA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*\*

1: Piri:\*\*  
2: Piri:\*\*  
3: Piri:\*\*  
4: Piri:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	955	97.9	576	1	RLCSD
2	865.5	88.8	564	1	RLCSAG
3	336	34.5	289	1	RLTGT
4	327.5	33.6	528	2	S32431
5	327.5	33.6	562	2	S16022
6	326	33.4	247	2	JU0393
7	326	33.4	247	2	JC5032
8	326	33.4	289	2	JC5006
9	323	33.1	527	2	S32430
10	310.5	31.8	251	2	C39761
11	303.5	31.1	528	1	TZLSA
12	300.5	30.8	278	2	S23519
13	293.5	30.1	250	2	JN0108
14	274.5	28.2	277	2	S22494
15	273	28.0	286	1	RLPUGG
16	273	28.0	570	2	S62627
17	272	27.9	286	2	S25560
18	270	27.7	245	2	JC4840
19	265	27.2	286	2	JC4235
20	264	27.1	254	2	PD0018
21	246.5	25.3	316	2	JT0753
22	186.5	19.1	294	2	S28421
23	174	17.8	278	2	A39817
24	170	17.4	313	2	S17757
25	168.5	17.3	261	2	JE0401
26	148.5	15.2	289	2	TJ2573
27	138	14.2	272	2	JC4811
28	131	13.4	253	2	S28542
29	127	13.0	253	2	S28539

## ALIGNMENTS

### RESULT 1

#### RLCSD

ricin D precursor - castor bean

N;Contains: rRNA N-glycosidase (EC 3.2.2.22)

C;Species: Ricinus communis (castor bean)

C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 16-Jul-1999

C;Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R;Hallig, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.

Nucleic Acids Res. 13, 8019-8033, 1985

A;Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A;Reference number: A24041; MUID:86067214; PMID:2399712

A;Accession: A24041

A;Molecule type: DNA

A;Residues: 1-576 <HAL>

A;Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R;Regear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A;Title: The lectin gene family of Ricinus communis: cloning of a functional ricin gene

A;Reference number: S20513; MUID:92163016; PMID:1371405

A;Accession: S20513

A;Molecule type: DNA

A;Residues: 1-576 <TRE>

A;Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R;Lamb, F.I.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A;Title: Nucleotide sequence of cloned cDNA coding for prepro-ricin.

A;Reference number: A24614; MUID:85179479; PMID:3838723

A;Accession: A24614

A;Molecule type: mRNA

A;Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>

A;Cross-references: GB:X03388; NID:g21077; PIDN:CAA26230.1; PID:g21078

R;Yoshitake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A;Title: Isolation and sequences of peptic peptides, and the complete sequence of Ile

A;Reference number: A03372

A;Accession: A03372

A;Molecule type: protein

A;Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 272-283, 'L', 285-288, 290-302 <YOS>

A;Note: this paper cites the others in the series providing experimental details for the

R;Araki, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1985

A;Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan

A;Reference number: A24010

A;Accession: A24010

A;Molecule type: protein

A;Residues: 315-383, 'PS', 386-576 <ARA>

R;Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A;Title: Primary structure of Ala chain of ricin D.

A;Reference number: A03374

A;Accession: A03374

A;Molecule type: protein

A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405-551, 'V', 553-557, 'E', 559-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F', 'FUN' <FUN>  
 A:Note: This paper, one of a series, summarizes the experimental details for the determination of the complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.  
 R:Ready, M.P.; Kim, Y.; Robertus, J.D.  
 Proteins 10, 270-278, 1991  
 A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism of ricin toxicity.  
 A:Reference number: A48233; PMID:91352006; PMID:1881883  
 A:Contents: annotation; active site  
 R:Rutenber, E.; Robertus, J.D.  
 Proteins 10, 260-269, 1991  
 A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.  
 A:Reference number: A48238; PMID:91352005; PMID:1881882  
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms  
 R:Katzin, B.J.; Collins, E.J.; Robertus, J.D.  
 Proteins 10, 251-259, 1991  
 A:Title: Structure of ricin A-chain at 2.5 angstroms.  
 A:Reference number: A48239; PMID:91352004; PMID:1881881  
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms  
 C:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which is inactive. The A chain inhibits protein synthesis; it inactivates the 60S ribosomal subunit of the cell of the A chain; B chains are also responsible for cell agglutination (lectin activity). This protein is cytotoxic and very poisonous to animals.  
 C:Superfamily: ricin; rRNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed  
 F:1-35/Domain: signal sequence #status predicted <SIG>  
 F:36-302/Product: ricin D chain A #status experimental <ACH>  
 F:46-293/Domain: rRNA N-glycosidase homology <RNG>  
 F:315-576/Product: ricin D chain B #status experimental <BCH>  
 F:331-373, 374-414, 417-455, 462-497, 501-540, 543-576/Region: 40-residue repeats  
 F:45, 409, 449/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F:115, 158, 243, 244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:212/Active site: Glu #status experimental  
 F:215/Active site: Arg #status predicted  
 F:294-318, 334-353, 377-394, 465-478, 504-521/Disulfide bonds: #status experimental  
 F:336, 349, 360/Binding site: N-acetylglucosamine (Asp, Gln, Asn) #status experimental  
 F:548, 569/Binding site: N-acetylglucosamine (Asp, Asn) #status experimental

Query Match 97.9%; Score 955; DB 1; Length 576;  
 Best Local Similarity 95.0%; Pred. No. 1.1e-79;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 2 IFPKQYPIINFITAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 36 IFPKQYPIINFITAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 95  
 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 111  
 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 155  
 112 GNYDRLEQLAGNLRNIELGNGLPEEALISALYYSTGGTQPLTARSFIIICQMISEAR 171  
 156 GNYDRLEQLAGNLRNIELGNGLPEEALISALYYSTGGTQPLTARSFIIICQMISEAR 215  
 172 FOYIEGEMTRIRYNRRA 190  
 216 FOYIEGEMTRIRYNRRA 234

RESULT 2  
 RLCSAG  
 agglutinin precursor - castor bean  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Ricinus communis (castor bean)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C:Accession: A24261; A24210  
 R:Roberts, L.M.; Lamb, I.F.; Pappin, D.J.C.; Lord, J.M.  
 J. Biol. Chem. 260, 15682-15686, 1985  
 A:Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.  
 A:Reference number: A24261; PMID:86059449; PMID:2999130  
 A:Accession: A24261  
 A:Molecule type: mRNA  
 A:Residues: 1-564 <ROB>  
 A:Cross-references: GB:M12089; NID:g169700; PID:AAA33869.1; PID:g169701

R:Araki, T.; Yoshioka, Y.; Funatsu, G.  
 Biochim. Biophys. Acta 872, 277-285, 1986  
 A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.  
 A:Reference number: A24210  
 A:Accession: A24210  
 A:Molecule type: protein  
 A:Residues: 303-325, 'F', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 553-557, 'E', 559-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F', 'FUN' <FUN>  
 C:Comment: This protein has strong agglutinating activity and weak cytotoxicity compared with ricin.  
 C:Superfamily: ricin; rRNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-290/Product: agglutinin chain A #status predicted <ACH>  
 F:55-281/Domain: rRNA N-glycosidase homology <RNG>  
 F:303-564/Product: agglutinin chain B #status experimental <BCH>  
 F:319-361, 362-402, 405-443, 450-485, 489-528, 531-564/Region: 40-residue repeats  
 F:34, 259/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:104, 147, 231, 232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:200, 203/Active site: Glu, Arg #status predicted  
 F:282-306, 322-341, 365-382, 453-466, 492-509/Disulfide bonds: #status predicted  
 F:324, 337, 348/Binding site: N-acetylglucosamine (Asp, Gln, Asn) #status predicted  
 F:397, 437/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F:536, 557/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 88.8%; Score 865.5; DB 1; Length 564;  
 Best Local Similarity 86.9%; Pred. No. 1.8e-71;  
 Matches 173; Conservative 7; Mismatches 8; Indels 11; Gaps 2;  
 2 IFPKQYPIINFITAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 25 IFPKQYPIINFITAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 111  
 85 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 144  
 112 GNYDRLEQLAGNLRNIELGNGLPEEALISALYYSTGGTQPLTARSFIIICQMISEAR 171  
 145 GNYDRLEQLAGNLRNIELGNGLPEEALISALYYSTGGTQPLTARSFIIICQMISEAR 203  
 172 FOYIEGEMTRIRYNRRA 190  
 204 FOYIEGEMTRIRYNRRA 222

RESULT 3  
 RLTSZ  
 rRNA N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolian snail  
 N:Alternate names: alpha-TCS; type I ribosome-inactivating protein  
 C:Species: Trichosanthin kirilowii (Mongolian snake-gourd)  
 C:Date: 30-Sep-1988 #sequence\_revision 26-Jan-1996 #text\_change 23-Mar-2001  
 C:Accession: J05566; A36274; J01093; A36273; J00003  
 R:Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.  
 Gene 97, 267-272, 1991  
 A:Title: Cloning of trichosanthin cDNA and its expression in Escherichia coli.  
 A:Reference number: J05566; PMID:91153657; PMID:1999291  
 A:Accession: J05566  
 A:Molecule type: mRNA  
 A:Residues: 1-289 <SHA>  
 A:Cross-references: GB:M34858; NID:g170536; PIDN:AAA34207.1; PID:g170537  
 A:Experimental source: tuber  
 R:Chow, T.P.; Feldman, R.A.; Lovett, M.; Piatak, M.  
 J. Biol. Chem. 265, 8670-8674, 1990  
 A:Title: Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I ribosome-inactivating protein.  
 A:Reference number: A36274; PMID:90256790; PMID:2341400  
 A:Accession: A36274  
 A:Molecule type: DNA  
 A:Residues: 1-233, 'T', 235-246, 'W', 248-289 <CHO>  
 A:Cross-references: GB:J05434; NID:g170534; PIDN:AAA34206.1; PID:g170535  
 R:Zheng, H.G.; Wang, B.; Shao, P.Z.; Yang, X.N.  
 Acta Genet. Sin. 21, 42-51, 1994  
 A:Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.  
 A:Reference number: J01093; PMID:94271613; PMID:8003348  
 A:Accession: J01093



A:Molecule type: DNA

A:Residues: 1-72, 'V', 74-90, 'S', 92-233, 'T', 235-267, 'D', 269-289 <ZHE>  
A:Cross-references: GB:S70176; NID:G547148; PIDN:AAB31048.1; PID:G547149  
R:Collins, E.J.; Robertus, J.D.; Lopresti, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwang, J. Biol. Chem. 265, 8665-8669, 1990

A:Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for abrin

A:Reference number: A36273; MUID:90256789; PMID:2341399

A:Accession: A36273

A:Molecule type: protein

A:Residues: 24-270 <COL>

R:Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z. Pure Appl. Chem. 58, 789-798, 1986

A:Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application

A:Reference number: JT0003

A:Accession: JT0003

A:Molecule type: protein

A:Residues: 24-56, 'L', 58-59, 'I', 61-71, 'I', 73-81, 85-86, 'L', 88-92, 'DAGLPENAVL', 93-142, 'GL'

A:Experimental source: tuber

R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

submitted to the Brookhaven Protein Data Bank, July 1994

A:Reference number: A67091; PDB:1MRJ

A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with adenine, residues 24-

R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

submitted to the Brookhaven Protein Data Bank, July 1994

A:Reference number: A67092; PDB:1MRK

A:Contents: annotation; X-ray crystallography, 1.6 angstroms, with formycin, residues 24-

R:Xiong, J.P.; Xia, Z.X.; Wang, Y.

submitted to the Brookhaven Protein Data Bank, December 1994

A:Reference number: A67111; PDB:1RCS

A:Contents: annotation; X-ray crystallography, 1.7 angstroms, with NADPH, residues 24-27

R:Xiong, J.P.; Xia, Z.X.; Wang, Y.

Nat. Struct. Biol. 1, 695-700, 1994

A:Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution re

A:Reference number: A58622; MUID:95360714; PMID:7634073

A:Contents: annotation; X-ray crystallography, 1.7 angstroms

A:Comment: Alpha-trichosanthin has been used to induce abortions.

C:Genetics:

A:Gene: tcs

C:Function:

A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then

C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology

C:Keywords: abortifacient; glycosidase; hydrolase; root; toxin

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-270/Product: trichosanthin alpha #status experimental <MAT>

F:27-266/Domain: rRNA N-glycosidase homology <RNG>

F:271-289/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:93,183,186/Active site: Tyr, Glu, Arg #status predicted

Query Match 34.5%; Score 336; DB 1; Length 289;

Best Local Similarity 38.9%; Pred. No. 2.8e-23;

Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 10 INFETAGATVQSTNFIHRAVRGLTVPN-----RVGLPINORFILVELSNHAE 59

Db 25 VSPRLSGATSSSYGVFISNLR---KALPNKLYDIPLLRSSLPQSRYALIHITNVADE 81

QY 60 SVTLALDVTNAYVGVYRAGNSAFVFFHDPNQEDA-EALTHLFTDVQNRVTFAGCNYDRLE 118

Db 82 TTSVALDVINVMVIMTRAGTSTFF---NEASATEAKYVFDAMRKVILPYSGNYERLQ 138

QY 119 QLAGNLRENIELNGPLEEASALYTYSTGTQTLPLARSFFIICQIMISEARFQYIEG 178

Db 139 TAAGKIRENIPGLPALDSAITLFFNYAN-----SAASALMVLIOSTSEARYKIEQ 193

QY 179 MRTRI 183

Db 194 IGRV 198

RESULT 4

S32431

abrin-d precursor - Indian licorice (fragment)

N;Contains: rRNA N-glycosidase (EC 3.2.2.22)

C;Species: Abrus precatorius (Indian licorice)

C;Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 01-Aug-1997

C;Accession: S32431; S34408

R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.

J. Mol. Biol. 229, 263-267, 1993

A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing. Cor

A:Reference number: S32429; MUID:93132798; PMID:8421313

A:Accession: S32431

A:Molecule type: mRNA

A:Residues: 1-528 <HUN>

A:Cross-references: GB:M98346

R:Hung, C.; Lee, M.; Lee, T.; Lin, J.

submitted to the EMBL Data Library, March 1993

A:Reference number: S34408

A:Accession: S34408

A:Molecule type: mRNA

A:Residues: 1-169, 'C', 171-320, 'L', 322-528 <HU2>

A:Cross-references: GB:M98346

C;Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating

The A and B chains are linked by a single disulfide bond, which is essential for toxicit

C:Superfamily: ricin; rRNA N-glycosidase homology

C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; p

F:1-251/Product: abrin-d chain A #status predicted <ACH>

F:7-246/Domain: rRNA N-glycosidase homology <RNG>

F:261-528/Product: abrin-d chain B #status predicted <BCH>

F:283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted

F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted

F:164,167/Active site: Glu, Arg #status predicted

F:200,253,361,401,402/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:247-269,286-305,323-346,417-430,456-473/Disulfide bonds: #status predicted

F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 33.6%; Score 327.5; DB 2; Length 528;

Best Local Similarity 43.9%; Pred. No. 3.6e-22;

Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;

QY 6 QVPINFTTAGATVQSTNFIHRAVRGLT-----VLPNRVGLPINORFILVELSNHA 57

Db 1 QQVVKFTTEGATSSQYKQFIEALRQLTGLIHDPVLPDPTTVEERNRYITVELSNSE 60

QY 58 ELSVTALDVTNAYVGVYRAGNSAFVFFHDPNQEDA-EALTHLFTDVQNRVTFAGCNYDRLE 117

Db 61 RESIEVGIDVTNAYVAVRAGSQSYFL---RDAPASASTYLPFGTQ-RYSLRFDGSGDL 116

QY 118 EQLAGNLRENIELNGPLEEASALYTYSTGTQTLPLARSFFIICQIMISEARFQYIEG 177

Db 117 ERWAHQTRREISLGLQALTHAIS---FLRSGASNDDEKARTLIVLIQMASEARVRYISN 173

QY 178 EMRTIR 184

Db 174 RVGVSR 180

RESULT 5

S16022

abrin-c precursor - Indian licorice

N;Contains: rRNA N-glycosidase (EC 3.2.2.22)

C;Species: Abrus precatorius (Indian licorice)

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999

C;Accession: S16022

R:Wood, K.A.; Lord, J.M.; Wawrzynczak, E.J.; Piatak, M.

Eur. J. Biochem. 198, 723-732, 1991

A:Title: Preproabrin: genomic cloning, characterisation and the expression of the A-chain

A:Reference number: S16022; MUID:91266957; PMID:2050149

A:Accession: S16022

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-562 <WOO>

A:Cross-references: EMBL:X55667; NID:g16084; PIDN:CAA39202.1; PID:g16085

C;Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating

The A and B chains are linked by a single disulfide bond, which is essential for toxicit

C:Superfamily: ricin; rRNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic acid  
 F:35-285/Product: abrin-c chain A #status predicted <ACH>  
 F:41-280/Domain: rRNA N-glycosidase homology <RNG>  
 F:295-369/Product: abrin-c chain B #status predicted <BCH>  
 F:317-359/Region: 40-residue repeats  
 F:335/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
 F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:198,201/Active site: Glu, Arg #status predicted  
 F:234,287,395,435,436/Binding site: carbohydride (Asn) (covalent) #status predicted  
 F:281-303,320-339,363-380,451-464,490-507/Disulfide bonds: #status predicted  
 F:322,346/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted  
 F:534,555/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 33.6%; Score 327.5; DB 2; Length 562;  
 Best Local Similarity 43.9%; Pred. No. 3.9e-22;  
 Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;

QY 6 QYPIINFTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILVELSNHA 57  
 Db 35 QDQIKFTFEGATSSQYKQFIALQRUTGGIHIPIVLPDPTTVEERNRYITVELSNSE 94  
 QY 58 ELSVTLALDVTNAYVYVGRAGNSAYFFHFPDNOEDAEATHLFTDVQNRYYTFAFGNRYDL 117  
 Db 95 RESIEVGIDVTNAYVYVGRAGNSQSYFL---RDAPASASTYLPFGTQ-RYSLRFDSYGD 150  
 QY 118 EQLAGNREINELGNPLEEASALYYSTGTQPTLARSFIIQIMISEARFOYIEG 177  
 Db 151 ERWQHQTREISLGLQALTHAIS---FLRSGANDEEKARTLIVIQWASEAARYIYSN 207  
 QY 178 EMRTIR 184  
 Db 208 RVGSIR 214

RESULT 6  
 JUU0393  
 Karasurin - Mongolian snake-gourd  
 C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
 C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 12-Apr-1995  
 C:Accession: JUU0393; PS0163  
 R:Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogiwara, Y.  
 Chem. Pharm. Bull. 39, 1244-1249, 1991  
 A:Title: The complete amino acid sequence of an abortifacient protein, karasurin.  
 A:Reference number: JUU0393; MUID:92005921; PMID:1914000  
 A:Accession: JUU0393  
 A:Molecule type: protein  
 A:Residues: 1-247 <TOY>  
 A:Note: a sequence which lacks Ala-247 is also shown in this publication  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 C:Keywords: abortifacient  
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.4%; Score 326; DB 2; Length 247;  
 Best Local Similarity 39.6%; Pred. No. 1.9e-22;  
 Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVR-----GRLTVLP-NRVGLPINQRFILVELSNHAEISVT 62  
 Db 2 VSRFLSGATSSSYGVFISNLKALPYERKLYDIPILRSTLPGSORYALIHITNYADETIS 61  
 QY 63 LALDVTNAYVYVGRAGNSAYFFHFPDNOEDA-EATHLFTDVQNRYYTFAFGNRYDL 121  
 Db 62 VAIDVTNAYVYVGRAGDSYFF---NEASATAAKYVFKDKRKVTLPSYGNVRLQIAA 118  
 QY 122 GNLRNIELGNPLEEASALYYSTGTQPTLARSFIIQIMISEARFOYIEGEMRT 181  
 Db 119 GKIRENIPGLPALDSAITTLFYNNAN-----SAASALMVLQISTSEARFYKFIQQIGK 173  
 QY 182 RI 183  
 Db 174 RV 175

RESULT 7  
 JCU5032  
 Karasurin-B - Trichosanthes kirilowii var. japonica  
 C:Species: Trichosanthes kirilowii var. japonica  
 C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997  
 C:Accession: JCU5032  
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996  
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka:  
 A:Reference number: JCU5032; MUID:97108848; PMID:8951169  
 A:Accession: JCU5032  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-247 <KON>  
 A:Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyti:  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.4%; Score 326; DB 2; Length 247;  
 Best Local Similarity 39.6%; Pred. No. 1.9e-22;  
 Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVR-----GRLTVLP-NRVGLPINQRFILVELSNHAEISVT 62  
 Db 2 VSRFLSGATSSSYGVFISNLKALPYERKLYDIPILRSTLPGSORYALIHITNYADETIS 61  
 QY 63 LALDVTNAYVYVGRAGNSAYFFHFPDNOEDA-EATHLFTDVQNRYYTFAFGNRYDL 121  
 Db 62 VAIDVTNAYVYVGRAGDSYFF---NEASATAAKYVFKDKRKVTLPSYGNVRLQIAA 118  
 QY 122 GNLRNIELGNPLEEASALYYSTGTQPTLARSFIIQIMISEARFOYIEGEMRT 181  
 Db 119 GKIRENIPGLPALDSAITTLFYNNAN-----SAASALMVLQISTSEARFYKFIQQIGK 173  
 QY 182 RI 183  
 Db 174 RV 175

RESULT 8  
 JCU5606  
 Karasurin C - Trichosanthes kirilowii var. japonica  
 N:Contains: karasurin A  
 C:Species: Trichosanthes kirilowii var. japonica  
 C:Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 19-Jul-2002  
 C:Accession: JCU5606; JCU5033  
 R:Mizukami, H.; Iida, K.; Kondo, T.; Ogiwara, Y.  
 Biol. Pharm. Bull. 20, 711-713, 1997  
 A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote:  
 A:Reference number: JCU5606; MUID:97356562; PMID:9212998  
 A:Accession: JCU5606  
 A:Molecule type: DNA  
 A:Residues: 1-289 <Miz>  
 A:Cross-references: DDBJ:AB000666; NID:G2329830; PIDN:BAA21786.1; PID:G2329831  
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996  
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka:  
 A:Reference number: JCU5032; MUID:97108848; PMID:8951169  
 A:Accession: JCU5033  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 22-270 <KON>  
 C:Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, aborti:  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:22-270/Product: karasurin C #status predicted <MAC>  
 F:24-270/Product: karasurin A #status predicted <MAA>  
 F:27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.4%; Score 326; DB 2; Length 289;  
 Best Local Similarity 39.6%; Pred. No. 2.3e-22;  
 Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

```

QY 10 INFATTAGTQSYNTNFIRAVR-----GRLTVLP-NRVGLPINQRFILVELSNHAEISVT 62
Db 25 VSRFLSGATSSYGVFTSNLRKALPYERKLYDIPLRSTLFGSQRYALIHLTNTVADETIS 84
QY 63 LALDVTNAVYVGYRAGNSAYFFHPDQDEA-EAITHLFTDVQNRVYTFAGGNYDRLEQLA 121
Db 85 VALDVTNAVYVGYRAGDTSYFF---NEASATEAAKYVFDKARKVTLPSYGNRYERLQATA 141
QY 122 GNLRNRELNGPLUEEAISALYYSTGTGTQTLPLARSFIICQMISEAARFYIEGEMRT 181
Db 142 GKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEAAKYKFFIEQIGK 196
QY 182 RI 183
Db 197 RV 198

RESULT 9
S32430
A:Title: precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 20-Aug-1999
C:Accession: S32430; JCI399
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A:Title: Primary structure of three distinct isoabrinins determined by cDNA sequencing.
A:Reference number: S32429; MUID:93132798; PMID:8421313
A:Accession: S32430
A:Molecule type: mRNA
A:Residues: 1-527 <HUN>
A:Cross-references: GB:M98345; NID:gl66296; PIDN:AAA32625.1; PID:gl66297.
R:Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A:Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic
A:Reference number: JCI398; MUID:93169023; PMID:7763422
A:Accession: JCI399
A:Molecule type: Protein
A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'M', 427, 'D', 429-430
A:Experimental source: Seed
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; P
F:1-250/Product: abrin-b chain A #status predicted <ACH>
F:7-245/Domain: rRNA N-glycosidase homology <RNG>
F:260-527/Product: abrin-b chain B #status experimental <BCH>
F:282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F:74,113,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:110,360,400/Binding site: carboxylate (Asn) (covalent) #status predicted
F:163,166/Active site: Glu, Arg #status predicted
F:246-268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted
F:287,311/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
F:499,520/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 33.1%; Score 323; DB 2; Length 527;
Best Local Similarity 43.9%; Pred. No. 9, 2e-22;
Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;

QY 6 QYPINFTAGATVQSYNTNFIRAVRGLT-----VLPNRVGLPINQRFILVELSNH 57
Db 1 QDOVIFKFTTEGATSQSYKQFTEALRQLRGLTHGIPVLPDPTTLQERNRYISVELNSD 60
QY 58 ELSVTLALDVTNAVYVGYRAGNSAYFFHPDQDEAETHLFTDVQNRVYTFAGGNYDR 117
Db 61 TESTEAGIDVSNVAVYVAVRAGNSRYFL-----RDAPTSASRYLFTGTQ-QYSLRFGSYIDL 116
QY 118 POLAGNLRNRELNGPLUEEAISALYYSTGTGTQTLPLARSFIICQMISEAARFYIEG 177
Db 117 ERLARQTFQQLPLGLQALRHAISEL-----QSGTDDQEFARTLVIQWASEAARYFLSY 172
QY 178 EMRTIRVYN 186

```

```

Db 173 RVGVSIPTN 181

RESULT 10
S39761
A:Title: precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C:Accession: C39761; S14471
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:91201329; PMID:2016300
A:Accession: C39761
A:Molecule type: DNA
A:Residues: 1-251 <EVE>
R:Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A:Description: Direct molecular cloning of two distinct abrin A-chains.
A:Reference number: S14471
A:Accession: S14471
A:Molecule type: DNA
A:Residues: 'M', 1-251 <EV2>
A:Cross-references: EMBL:X54872; NID:gl6088; PIDN:CAA38654.1; PID:gl6089
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin
F:1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>
F:7-246/Domain: rRNA N-glycosidase homology <RNG>
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted

Query Match 31.8%; Score 310.5; DB 2; Length 251;
Best Local Similarity 43.2%; Pred. No. 5e-21;
Matches 79; Conservative 22; Mismatches 67; Indels 15; Gaps 4;

QY 10 INFATTAGTQSYNTNFIRAVRGL-----TVLPNRVGLPINQRFILVELSNHAEISV 61
Db 5 IKFTEGATSQSYKQFTEALRQLRGLTHGIPVLRDPTTVEERNRYITVELSNSRESI 64
QY 62 TLALDVTNAVYVGYRAGNSAYFFHPDQDEAETHLFTDVQNRVYTFAGGNYDRLEQLA 121
Db 65 EVGIDVTNAVYVAVRAGSQSYFL---RDAPASATYLTGTQ-RYSRFDGSGYDLERWA 120
QY 122 GNLRNRELNGPLUEEAISALYYSTGTGTQTLPLARSFIICQMISEAARFYIEGEMRT 181
Db 121 HQTREQISLGLQALTHAIS---FLRSGASNDKARTLVIQWASEAARYISNRGV 177
QY 182 RIR 184
Db 178 SIR 180

RESULT 11
TZLSA
A:Title: precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 31-Dec-1993 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999
C:Accession: S32429; JCI398; S14472; S24133; S74110; S74111
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A:Title: Primary structure of three distinct isoabrinins determined by cDNA sequencing.
A:Reference number: S32429; MUID:93132798; PMID:8421313
A:Accession: S32429
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'E', 2-528 <HUN>
A:Cross-references: GB:M98344; NID:gl66294; PIDN:AAA32624.1; PID:gl66295.
A:Note: the coding region for the sequence shown is preceded by an ATG codon
R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Yanaka, M.
Agric. Biol. Chem. 52, 1095-1097, 1988

```

QY 10 INFTTAGATVQSYYTNFIRAVRGRL-----TVLPNRVGLPINQRFILVELSNHAELSV 61  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 5 IKFSTEGATSQSKQFIKALRERLGGLLHDIPVLDPDPTTLQERNRYITVELSNSDTESI 64  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 62 TLALDVNTNAYVGYRAGNSAYFFH--PNQEDAEATHLFTDVQNRYTFAPGCVDRLEQ 119  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 65 EYGVIDVTNAYWAYVAGTQSYFLRDAPSASD-----YLFTGT-DOHSLPFFTYGTDLER 118  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 120 LAGNLRENIELNGPLGLEBAISALYYVYSTGCTQLPTLARSFFICMIOMISEAARFOYIEGSM 179  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 119 WAHQSRQQIPIGLQALTHCIS---PFRSGNDNEKARTLIIVIQWVAEAPRTYSNRV 175  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 180 RTRIR 184  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 176 RVSIQ 180  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
  
RESULT 12  
S23519  
beta-luffin - smooth loofah  
C;Species: Luffa cylindrica (smooth loofah)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Aug-1999  
C;Accession: S23519; S23113  
R;Katoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koiwai, A.  
Plant Mol. Biol. 19, 887-889, 1992  
A;Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-inactivating  
A;Reference number: S23519; PMID:92353400; PMID:1643290  
A;Accession: S23519  
A;Molecule type: mRNA  
A;Residues: 1-278 <CAT>  
A;Cross-references: EMBL:X62372; NID:g19149; PIDN:CRA44230.1; PID:g19150  
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F;26-264/Domain: rRNA N-glycosidase homology <RNG>  
  
Query Match 30.8%; Score 300.5; DB 2; Length 278;  
Best Local Similarity 35.3%; Pred. No. 4.7e-20;  
Matches 66; Conservative 43; Mismatches 57; Indels 21; Gaps 4;  
  
QY 10 INFTTAGATVQSYYTNFIRAVRGRL-----TVLPNRVGLPINQRFILVELSNHAEL 59  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 24 VSFSLSGADSKSVKFPITALRKALPSKEKVSNIPLLPSAGA---SRYLMLQLSNYDAK 80  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 60 SYTLALDVNTNAYVGYRAGNSAYFFHPDNQEDAEATHLFTDVQNRYTFAPGCVDRLEQ 119  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 81 AITMAIDVTNVIMGYLVNSTSYFF---NESDAKLASQYVFKGSTIVTLPGSNVERLQN 137  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 120 LAGNLRENIELNGPLGLEBAISALYYVYSTGCTQLPTLARSFFICMIOMISEAARFOYIEGSM 179  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 138 AAGKYREKIPLGFRAFDSAITSLFHYDS-----TAAGAFVIIQTAAASFKEIGOI 192  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 180 RTRIRYN 186  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 193 IERI PKN 199  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
  
RESULT 13  
JN0108  
luffin-b - smooth loofah  
C;Species: Luffa cylindrica (smooth loofah)  
C;Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 07-May-1999  
C;Accession: JN0108  
R;Iselam, M.R.; Hirayama, H.; Funatsu, G.  
Agric. Biol. Chem. 55, 229-238, 1991  
A;Title: Complete amino acid sequence of luffin-b, a ribosome-inactivating protein from  
A;Reference number: JN0108; PMID:91248488; PMID:1368666  
A;Accession: JN0108  
A>Status: preliminary  
A;Molecule type: Protein  
A;Residues: 1-250 <ISL>  
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F;5-246/Domain: rRNA N-glycosidase homology <RNG>



F;24-269/Product: rRNA N-glycosidase alpha-monomer charin #status experimental <MAT>  
F;27-266/Domain: rRNA N-glycosidase homology <RNG>  
F;270-286/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F;93,183,186/Active site: Tyr, Glu, Arg #status predicted  
F;250/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 28.0%; Score 273; DB 1; Length 286;  
Best Local Similarity 34.8%; Pred. No. 1.6e-17;  
Matches 63; Conservative 40; Mismatches 62; Indels 16; Gaps 4;

```
QY 10 INFTTAGATVQSYTNFIRAVRGRLTVLPNRVGLPI-----NORFILVELSNHAELSVT 62
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 25 VSFRLSGADPRSYGNFIKDLRNALPFREKYNIPILLPSVSGAGRYLLMHLFNYDGKTIIT 84
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 63 LALDVTNAYVVGVRAGNSGAYFFHPDNQEDAE-AITHLFTDVQNRVTFAPGGNYDRLEQLA 121
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 85 VAVDVTVNVYIMGYLADTTSYFF--NEPAAELASQYVFRDARRKKTLPYSGNYERLQIAA 141
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 GNLRENIELGNGPLEEAISALVYVYSTGGTQLPFLARSFLICIMISEAARFOYIEGEMRT 181
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 142 GKPREKIPIGLPALDSASTLLHYDS-----TAAAGALLVLIQTAAEARFKYIEQIOE 196
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 182 R 182
Db 197 R 197
```

Search completed: February 10, 2004, 16:28:04  
Job time : 11.2817 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:55 ; Search time 6.3803 Seconds  
(without alignments)  
1409.756 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MIPKQYPIINFITAGATVQ.....RFQVIEGEMTRIRNRRSA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	955	97.9	576	1	RICI_RICCO
2	865.5	88.8	564	1	AGGL_RICCO
3	336	34.5	289	1	RIP1_TRIKI
4	327.5	33.6	562	1	ABRC_ABRPR
5	326	33.4	289	1	RIPS_TRIKI
6	323	33.1	527	1	ABRB_ABRPR
7	320.5	32.9	282	1	RIP2_BRYDI
8	307.5	31.5	563	1	NIGB_SAMNI
9	305	31.3	290	1	RIP1_BRYDI
10	303.5	31.1	528	1	ABRA_ABRPR
11	293.5	30.1	250	1	RIPB_LUFYC
12	281.5	28.9	286	1	RIP1_CUCFI
13	274.5	28.2	277	1	RIPA_LUFYC
14	273	28.0	286	1	RIP1_MOMCH
15	272	27.9	286	1	RIP2_MOMBA
16	264	27.1	254	1	MLA_VISAL
17	257	26.4	294	1	RIP1_TRIAN
18	246.5	25.3	316	1	RIPG_GELMU
19	186.5	19.1	294	1	RIPA_PHYAM
20	170	17.8	278	1	RIPP_MIRVA
21	170	17.4	313	1	RIP1_PHYAM
22	168.5	17.3	261	1	RIPS_PHYAM
23	131	13.4	253	1	RIP7_SAPOF
24	127	13.0	253	1	RIP5_SAPOF
25	125	12.8	310	1	RIP2_PHYAM
26	124	12.7	299	1	RIP6_PHYAM
27	121	12.4	292	1	RIP2_SAPOF
28	116.5	11.9	293	1	RIP0_DIAAC
29	114	11.7	280	1	RIP2_HORVU
30	110	11.3	236	1	RIP3_SAPOF
31	108	11.1	319	1	SLTA_BP933
32	107	11.0	280	1	RIP1_HORVU
33	103.5	10.6	300	1	RIP3_MAIZE

## RESULT 1

ID	RICI_RICCO	STANDARD;	PRT;	576 AA.
AC	P02879; P02880;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Ricin B chain].			
OS	Ricinus communis (Castor bean).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.			
OX	NCBI_TaxID=3988;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86067214; PubMed=2999712;			
RA	Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L., Weaver R.F.;			
RA	"Genomic cloning and characterization of a ricin gene from Ricinus communis.";			
RT	comminis.";			
RL	Nucleic Acids Res. 13:8019-8033(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92163016; PubMed=1371405;			
RA	Tregear J.W., Roberts L.M.;			
RT	"The lectin gene family of Ricinus communis: cloning of a functional ricin gene and three lectin pseudogenes.";			
RL	Plant Mol. Biol. 18:515-525(1992).			
RN	[3]			
RP	SEQUENCE OF 12-576 FROM N.A.			
RX	MEDLINE=85179479; PubMed=3838723;			
RA	Lamb A., Roberts L.M., Lord J.M.;			
RT	"Nucleotide sequence of cloned cDNA coding for preproricin.";			
RL	Eur. J. Biochem. 148:265-270(1985).			
RN	[4]			
RP	SEQUENCE OF 36-302.			
RX	Yoshitake S., Funatsu G., Funatsu M.;			
RT	"Isolation and sequences of peptic peptides, and the complete sequence of Ile chain of ricin-D.";			
RL	Agric. Biol. Chem. 42:1267-1274(1978).			
RN	[5]			
RP	SEQUENCE OF 315-576.			
RX	Funatsu G., Kimura M., Funatsu M.;			
RT	"Primary structure of Ala chain of ricin D.";			
RL	Agric. Biol. Chem. 43:2221-2224(1979).			
RN	[6]			
RP	CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.			
RX	MEDLINE=90344223; PubMed=1368517;			
RA	Kimura Y., Kusuku H., Tada M., Takagi S., Funatsu G.;			
RT	"Structural analyses of sugar chains from ricin A-chain variant.";			
RL	Agric. Biol. Chem. 54:157-162(1990).			
RN	[7]			
RP	REVIEW.			
RX	MEDLINE=21480122; PubMed=11595634;			
RA	Olanes S., Kozlov J.V.;			

P28522 zea mays (m)  
P25892 zea mays (m)  
P08026 bacterioph  
P10149 bacterioph  
O67411 aquifex ae  
Q00531 hordeum vul  
P75255 mycoplasma  
Q95xb4 homo sapien  
P24153 vibrio chol  
P27561 saponaria o  
Q00971 vibrio prot  
P44755 haemophilus

## ALIGNMENTS



RT "Ricin.";

RL Toxicon 39:1723-1728 (2001).

RP [8]

RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).

RA MEDLINE=87165983; PubMed=3558397;

RA Monfort W., Villafraanca J.E., Monzingo A.F., Ernst S.R., Katzin B.,

RA Rutenber E., Xiong N.H., Hamlin R., Robertus J.D.;

RT "The three-dimensional structure of ricin at 2.8 A.";

RL J. Biol. Chem. 262:5398-5403 (1987).

RL [9]

RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.

RX MEDLINE=91352004; PubMed=1881881;

RA Katzin B.J., Collins E.J., Robertus J.D.;

RT "Structure of ricin A-chain at 2.5 A.";

RL Proteins 10:251-259 (1991).

RL [10]

RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.

RX MEDLINE=91352005; PubMed=1881882;

RA Rutenber E., Robertus J.D.;

RT "Structure of ricin B-chain at 2.5-A resolution.";

RL Proteins 10:260-269 (1991).

RL [11]

RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.

RX MEDLINE=95082010; PubMed=7990130;

RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,

RA Paupit R.A.;

RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";

RL J. Mol. Biol. 244:410-422 (1994).

RL [12]

RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.

RX MEDLINE=96374222; PubMed=8780513;

RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,

RA Molina-Svinth M.C., Robertus J.D.;

RT "Structure and activity of an active site substitution of ricin A chain.";

RL Biochemistry 35:11098-11103 (1996).

RL [13]

RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.

RX MEDLINE=97240820; PubMed=9086280;

RA Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Milne G.W.,

RA Robertus J.D.;

RT "Structure-based identification of a ricin inhibitor.";

RL J. Mol. Biol. 266:1043-1049 (1997).

RL [14]

RP MUTAGENESIS.

RX MEDLINE=93165632; PubMed=1287657;

RA Kin Y., Robertus J.D.;

RT "Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography.";

RL Protein Eng. 5:775-779 (1992).

CC -I- FUNCTION: Ricin is highly toxic to animal cells and to a less extent to plant cells. The A chain is responsible for inhibiting protein synthesis through the catalytic inactivation of 60S ribosomal subunits. It acts as a glycosidase that removes a specific adenine residue from an exposed loop of 28S ribosomal RNA. As this loop is involved in the binding of elongation factors, the modified ribosomes are unable to support protein synthesis. The A chain can inactivate a few thousand ribosomes per minute, thus inactivating them faster than the cell can make new ones. A single A-chain molecule can therefore kill an animal cell. The B chain binds to cell receptors and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (lectin activity). It binds to beta-D-galactopyranoside moieties.

CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.

CC -I- SUBUNIT: Disulfide-linked dimer of A and B chains.

CC -I- DOMAIN: The B chain is composed of two domains, each domain consists of 3 homologous subdomains (alpha, beta, gamma).

CC -I- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.

CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

CC -I- SIMILARITY: Contains 2 ricin B-type lectin domains.

CC -I- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).

CC -I- DATABASE: NAME=Protein Spotlight;

CC NOTE=Issue 31 of February 2003;

CC WWW="http://www.expasy.org/spotlight/articles/sptlt031.html".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X03179; CAA26939.1; -

DR EMBL; X52908; CAA37095.1; -

DR EMBL; X02388; CAA26230.1; -

DR EMBL; A12892; CAA01058.1; -

DR FIR; A24041; RUCSD.

DR PDB; 2AAI; 31-JAN-94.

DR PDB; 1APG; 31-JAN-94.

DR PDB; 1FMP; 31-OCT-93.

DR PDB; 1IFS; 14-JAN-98.

DR PDB; 1IFT; 14-JAN-98.

DR PDB; 1IFU; 14-JAN-98.

DR PDB; 1RTC; 31-OCT-93.

DR PDB; 1OBS; 16-JUN-97.

DR PDB; 1OBT; 16-JUN-97.

DR PDB; 1BR5; 02-SEP-98.

DR PDB; 1BR6; 02-SEP-98.

DR PDB; 1IL3; 16-JAN-02.

DR PDB; 1IL4; 16-JAN-02.

DR PDB; 1IL9; 16-JAN-02.

DR GlycoSuiteDB; P02879; -

DR InterPro; IPR00772; Ricin\_B\_lectin.

DR InterPro; IPR001574; RIP.

DR Pfam; PF00852; Ricin\_B\_lectin; 6.

DR Pfam; PF00161; RIP; 1.

DR PRINTS; PR00396; SHIGARICIN.

DR SMART; SM00458; RICIN; 2.

DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.

DR PROSITE; PS00275; SHIGA\_RICIN; 1.

KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;

KW Glycoprotein; Lectin; Signal; 3D-structure.

FT SIGNAL 1 35

FT CHAIN 36 302

FT PEPTIDE 303 314

FT CHAIN 315 576

FT DOMAIN 321 448

FT DOMAIN 451 575

FT REPEAT 331 373

FT REPEAT 374 414

FT REPEAT 417 449

FT REPEAT 462 497

FT REPEAT 501 540

FT REPEAT 543 570

FT ACT\_SITE 212 212

FT DISULFID 294 318

FT DISULFID 334 353

FT DISULFID 377 394

FT DISULFID 465 478

FT DISULFID 504 521

FT CARBOHYD 45 45

FT CARBOHYD 271 271

FT CARBOHYD 409 409

FT CARBOHYD 449 449

FT CARBOHYD 76 76

FT CONFLICT 551 551

FT STRAND 43 47

FT TURN 49 50

FT N-LINKED (GLCNAC. . .).

FT /FTIC-CAR 000080.

FT N-LINKED (GLCNAC. . .) (IN MINOR FORM).

FT /FTIC-CAR 000081.

FT N-LINKED (GLCNAC. . .).

FT N-LINKED (GLCNAC. . .).

FT E -> D (IN REF. 3).

FT CONFLICT 551 551

FT STRAND 43 47

FT TURN 49 50



Query Match 97.8%; Score 955; DB 1; Length 576;  
 Best Local Similarity 95.0%; Pred.No. 1.5e-81;  
 Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLLT-----VLPNRVGLPINQRFILV 51  
 DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLLTGDVRRHEIPVLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 111  
 DB 96 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRNIELGNGLPELEAISALYYSTGGTQLPPLARSFIIQIMISEAAR 171  
 DB 156 GNYDRLEQLAGNLRNIELGNGLPELEAISALYYSTGGTQLPPLARSFIIQIMISEAAR 215  
 QY 172 FOYIEGEMTRIRYNRRA 190  
 DB 216 FOYIEGEMTRIRYNRRA 234

## RESULT 2

AGGL\_RICCO STANDARD; PRT; 564 AA.  
 AC P06750;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
 OS Ricinus communis (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 OX NCBI\_TaxID=3988;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=86059449; PubMed=2999130;  
 RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;  
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";  
 RL J. Biol. Chem. 260:15682-15686(1985).  
 RN [2]  
 RN SEQUENCE OF 303-564.  
 RP TISSUE=Seed.  
 RC Araki T., Yoshioka Y., Funatsu G.;  
 RA "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";  
 RL Biochim. Biophys. Acta 872:277-285(1986).  
 RN [3]

SEQUENCE OF 303-337.  
 RX MEDLINE=80178723; PubMed=6768555;  
 RA Lin T.-S., Li S.-L.;  
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";  
 RL Eur. J. Biochem. 105:453-459(1980).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M12089; AAA33869.1; -;  
 CC EMBL; S40368; AAB22584.1; -;

DR PIR; A24261; RLCSAG.  
 DR HSP; P02879; IIR6.  
 DR Glycosylated; P06750; -;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP\_1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS0275; SHIGA\_RICIN; 1.  
 DR Plant defense; Hydroxylase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal; POTENTIAL.  
 FT SIGNAL 1 24  
 FT CHAIN 25 290 AGGLUTININ A CHAIN.  
 FT PROPEP 291 302 LINKER PEPTIDE.  
 FT CHAIN 303 564 AGGLUTININ B CHAIN.  
 FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 319 361 1-ALPHA.  
 FT REPEAT 362 402 1-BETA.  
 FT REPEAT 405 437 1-GAMMA.  
 FT REPEAT 450 485 2-ALPHA.  
 FT REPEAT 489 528 2-BETA.  
 FT REPEAT 531 558 2-GAMMA.  
 FT ACT\_SITE 200 200 BY SIMILARITY.  
 FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 322 341 BY SIMILARITY.  
 FT DISULFID 365 382 BY SIMILARITY.  
 FT DISULFID 453 466 BY SIMILARITY.  
 FT DISULFID 492 509 BY SIMILARITY.  
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .).  
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. .).  
 FT CONFLICT 331 331 F -> T (IN REF. 2).  
 FT CONFLICT 362 362 N -> D (IN REF. 2).  
 FT CONFLICT 374 374 R -> G (IN REF. 2).  
 FT CONFLICT 404 404 R -> T (IN REF. 2).  
 FT CONFLICT 552 552 F -> V (IN REF. 2).  
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;  
 Query Match 88.8%; Score 865.5; DB 1; Length 564;  
 Best Local Similarity 86.9%; Pred.No. 3.3e-73;  
 Matches 173; Conservative 7; Mismatches 8; Indels 11; Gaps 2;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLLT-----VLPNRVGLPINQRFILV 51  
 DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVGRLLTGDVRRHEIPVLPNRVGLPINQRFILV 84  
 QY 52 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 111  
 DB 95 ELSNHAELSVTLALDVNTNAVYVGRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 144  
 QY 112 GNYDRLEQLAGNLRNIELGNGLPELEAISALYYSTGGTQLPPLARSFIIQIMISEAAR 171  
 DB 145 GNYDRLEQLAGNLRNIELGNGLPELEAISALYYSTGGTQLPPLARSFIIQIMISEAAR 203  
 QY 172 FOYIEGEMTRIRYNRRA 190  
 DB 204 FOYIEGEMTRIRYNRRA 222  
 RESULT 3  
 ID RIPT\_TRIKI STANDARD; PRT; 289 AA.  
 AC P0989;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ribosome-inactivating protein alpha-trichosanthin precursor (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.

OX NCBI\_TaxID=3677;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=Maximowicz;

RC MEDLINE=91153657; PubMed=199291;

RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;

RT "Cloning of trichosanthin cDNA and its expression in *Escherichia*

RT *coli*."

RL Gene 97:267-272(1991).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Maximowicz; TISSUE=Leaf;

RX MEDLINE=90256790; PubMed=2341400;

RA Chow T., Feldman R.A., Lovett M., Piatak M.;

RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a

RT type I ribosome-inactivating protein."

RL J. Biol. Chem. 265:8670-8674(1990).

RN [3]

RP SEQUENCE OF 24-270.

RC STRAIN=Maximowicz; TISSUE=Tuberous root;

RX MEDLINE=90256789; PubMed=2341399;

RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,

RT "Primary amino acid sequence of alpha-trichosanthin and molecular

RT models for abrin A-chain and alpha-trichosanthin."

RL J. Biol. Chem. 265:8665-8669(1990).

RN [4]

RP SEQUENCE OF 24-270.

RC TISSUE=tuberous root;

RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,

RA Tian G.Y., Ni C.Z.;

RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and

RT application."

RL Pure Appl. Chem. 58:789-798(1986).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).

RX MEDLINE=94344957; PubMed=8086085;

RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;

RT "Structure of trichosanthin at 1.88-A resolution."

RL Proteins 19:4-13(1994).

RN [6]

RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).

RX MEDLINE=95344383; PubMed=7619070;

RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;

RT "Studies on crystal structures, active-centre geometry and

RT depurinating mechanism of two ribosome-inactivating proteins."

RL Biochem. J. 309:285-298(1995).

CC [1-]

CC FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS

CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT

CC INACTIVATES ENKARYOTIC 60S RIBOSOMAL SUBUNITS.

CC CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one

CC specific adenosine on the 28S rRNA.

CC [1-] SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

CC TYPE 1 RIP SUPERFAMILY.

CC [1-]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC [1-]

CC EMBL; M34858; AAA34207.1; -.

CC EMBL; J05434; AAA34206.1; -.

CC PIR; JT0566; RLTZT

CC PDB; 1MRJ; 07-FEB-95.

CC PDB; 1MRK; 07-FEB-95.

CC PDB; 1TCS; 10-JUL-95.

CC PDB; 1J4G; 28-JAN-03.

DR

DR PDB; INLI; 21-JAN-03.  
 DR PDB; IQD2; 24-APR-00.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PRO0396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 270

FT PROPEP 271 289  
 FT ACT SITE 183  
 FT CONFLICT 57 60  
 FT CONFLICT 82 84  
 FT CONFLICT 87 87  
 FT CONFLICT 92 92  
 FT CONFLICT 143 144  
 FT CONFLICT 196 196  
 FT CONFLICT 215 216  
 FT CONFLICT 231 231  
 FT CONFLICT 234 234  
 FT CONFLICT 246 266  
 FT CONFLICT 247 247  
 FT STRAND 25 28  
 FT TURN 30 31  
 FT HELIX 34 46  
 FT TURN 47 47  
 FT STRAND 50 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 89 92  
 FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT TURN 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT TURN 142 142  
 FT STRAND 145 147  
 FT STRAND 150 150  
 FT HELIX 152 163  
 FT TURN 164 165  
 FT HELIX 167 180  
 FT TURN 181 181  
 FT HELIX 182 186  
 FT STRAND 187 187  
 FT HELIX 188 195  
 FT TURN 196 196  
 FT STRAND 202 202  
 FT HELIX 206 226  
 FT TURN 227 230  
 FT STRAND 231 239  
 FT TURN 241 242  
 FT STRAND 245 250  
 FT TURN 251 252  
 FT HELIX 254 258  
 FT TURN 259 259  
 FT STRAND 260 260  
 FT STRAND 263 263  
 FT TURN 266 268

FT SEQUENCE 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;  
 TRICHOSANTHIN.  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q -> T (IN REF. 4).  
 S -> T (IN REF. 2).  
 MISSING (IN REF. 4).  
 T -> M (IN REF. 2).

RIBOSOME-INACTIVATING PROTEIN ALPHA-  
 MISSING IN MATURE PROTEIN.  
 BY SIMILARITY.  
 IPLL -> LPLI (IN REF. 4).  
 MISSING (IN REF. 4).  
 I -> L (IN REF. 4).  
 V -> VDAGLPNAVLI (IN REF. 4).  
 KI -> GL (IN REF. 4).  
 K -> S (IN REF. 4).  
 WS -> LWL (IN REF. 4).  
 Q ->





FT CONFLICT 291 D -> N (IN REF. 2).  
 FT CONFLICT 350 AE -> PQ (IN REF. 2).  
 FT CONFLICT 378 S -> N (IN REF. 2).  
 FT CONFLICT 426 L -> M (IN REF. 2).  
 FT CONFLICT 428 Y -> D (IN REF. 2).  
 FT CONFLICT 431 N -> S (IN REF. 2).  
 FT CONFLICT 431 R -> K (IN REF. 2).  
 FT CONFLICT 484 R -> S (IN REF. 2).  
 FT CONFLICT 491 N -> Y (IN REF. 2).  
 FT CONFLICT 493 H -> G (IN REF. 2).  
 FT CONFLICT 502 E -> Q (IN REF. 2).  
 FT CONFLICT 509 E -> Q (IN REF. 2).  
 FT CONFLICT 513 H -> W (IN REF. 2).  
 FT CONFLICT 516 H -> T (IN REF. 2).  
 SQ SEQUENCE 527 AA; 3253AE490CE9494A CRC64;  
 Query Match 33.1%; Score 323; DB 1; Length 527;  
 Best Local Similarity 43.9%; Pred. No. 1.3e-22;  
 Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;  
 QY 6 QYPLINFTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFILVELSNHA 57  
 Db 1 QDQVKEITTEGASQSYKQFIEALRQLTGGLIHGIPVLPDPTTLOERNYISVELNSND 60  
 QY 58 ELSVTALDVTNAYVGYRAGNSAYFFHPDQDAEAITHLFTDVQVRYTFAPGNYDRL 117  
 Db 61 TESTEAGIDVSNVAVYRAGNRSYFL---RDAPTSASRYLFTGTQ-QYSLRFNGSYIDL 116  
 QY 118 EOLAGNLENIELGNGLPLEEASLYYSTGGTQLPPLARSFFLICOMISEARFOYIE 177  
 Db 117 ERLARQTRQPLGLQALRHAIISFL-----QSGTDDQEIARTLVIHQWASEARYRFSY 172  
 QY 178 EMRTIRYN 186  
 Db 173 RVGSIRTN 181  
 RESULT 7  
 ID RIP2 BRYDI STANDARD; PRT; 282 AA.  
 AC P98184; Q98J0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 18-OCT-2001 (Rel. 40, Last sequence update)  
 DE Ribosome-inactivating protein bryodin II precursor (rRNA N-  
 Glycosidase) (EC 3.2.2.22) (BD2).  
 OS Bryonia dioica (Red bryony).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosoids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 OX NCBI\_TaxID=3652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Slegall C.B., Gawlak S.L., Marquardt H.;  
 RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant  
 Bryonia dioica";  
 RL Patent number US5597569, 28-JAN-1997.  
 RN [2]  
 RP SEQUENCE OF 22-42.  
 RC TISSUE=Root;  
 RX MEDLINE=95151812; PubMed=7849072;  
 RA Slegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
 RA Marquardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from  
 Bryonia dioica and their utility as carcinoma-reactive  
 immunocjugates";  
 RL Bioconj. Chem. 5:423-429(1994).  
 CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; I34238; ; NOT\_ANNOTATED\_CDS.  
 DR HSSP; P09989; IMRJ.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KM Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
 KM Multigene family; Glycoprotein; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 282  
 FT ACT SITE 183 183  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;  
 Query Match 32.9%; Score 320.5; DB 1; Length 282;  
 Best Local Similarity 46.0%; Pred. No. 1e-22;  
 Matches 81; Conservative 24; Mismatches 54; Indels 17; Gaps 7;  
 QY 10 INFTTAGATVQSYTNFIRAVRGRLT-LPNRVGLPINQ-----BFLVLSNHAELSV 61  
 Db 24 INFSLGATGATYKTRIRNLTGTVGTPRVYDIPVRNAAAGLARFQVLTNTNGESV 83  
 QY 62 TLALDVTNAYVGYRAGNSAYFFHPDQDAEAITHLFTDVQVRYTFAPGNYDRLQLA 121  
 Db 84 TVALDVTNAYVGYRAGNSAYFFHPDQDAEAITHLFTDVQVRYTFAPGNYDRLQLA 139  
 QY 122 GNL-RENIELGNGLPLEEASLYYSTGGTQLPPLARSFFLICOMISEARFOYIE 176  
 Db 140 GRISRENIELGFSEISSAIGNMFRNP-GTSVP---RAFTVIQTVSEARFKYIE 191  
 RESULT 8  
 ID NIGB SAMNI STANDARD; PRT; 563 AA.  
 AC P33183; P33184; P93542;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Nigrin b precursor (Agglutinin V) (SNAV) [Contains: Nigrin b A chain  
 (rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain].  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven E., Peumans W.J.;  
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V  
 (nigrin b), a GalNAc-specific type-2 ribosome-inactivating protein  
 from the bark of elderberry (Sambucus nigra)";  
 RL Eur. J. Biochem. 237:505-513(1996).  
 RN [2]  
 RP SEQUENCE OF 26-49 AND 298-321.  
 RC TISSUE=Bark;  
 RX MEDLINE=96215449; PubMed=8647092;  
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven E., Peumans W.J.;  
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V  
 (nigrin b), a GalNAc-specific type-2 ribosome-inactivating protein  
 from the bark of elderberry (Sambucus nigra)";  
 RL Eur. J. Biochem. 237:505-513(1996).  
 RN [2]  
 RP SEQUENCE OF 26-49 AND 298-321.  
 RC TISSUE=Bark;  
 RX MEDLINE=94003077; PubMed=8400135;  
 RA Girbes T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R.,  
 RA Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;  
 RT "Isolation and partial characterization of nigrin b, a non-toxic  
 novel type 2 ribosome-inactivating protein from the bark of Sambucus  
 nigra L.";  
 RL Plant Mol. Biol. 22:1181-1186(1993).  
 CC -!- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN



or send an email to license@isb-sib.ch).

CC EMBL: 124020; -- NOT\_ANNOTATED\_CDS.  
DR PIR: S16491; S16491.  
DR FDB: 1BR1; 04-MAR-98.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
KW 3D-structure; Multigene family; Glycoprotein; Signal.  
FT SIGNAL 1 23  
FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.  
FT PROPEP 271 290 MISSING IN MATURE PROTEIN.  
FT ACT\_SITE 183 183 BY SIMILARITY.  
FT ACT\_SITE 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT MUTAGEN 212 212 E->K: REDUCES ACTIVITY 10-FOLD.  
FT CONFLICT 61 65 RSSIS -> LRHXI (IN REF. 3).  
FT STRAND 25 28  
FT TURN 30 31  
FT TURN 34 34  
FT HELIX 47 47  
FT TURN 47 47  
FT STRAND 50 54  
FT TURN 55 56  
FT STRAND 57 60  
FT HELIX 66 69  
FT STRAND 70 76  
FT TURN 78 79  
FT STRAND 82 88  
FT TURN 89 92  
FT STRAND 93 99  
FT TURN 100 101  
FT STRAND 102 105  
FT HELIX 109 114  
FT TURN 115 117  
FT TURN 120 121  
FT STRAND 124 127  
FT HELIX 134 141  
FT TURN 142 142  
FT HELIX 145 147  
FT STRAND 150 150  
FT TURN 164 165  
FT HELIX 167 186  
FT STRAND 187 187  
FT HELIX 188 196  
FT STRAND 202 202  
FT HELIX 206 213  
FT TURN 214 214  
FT HELIX 215 225  
FT TURN 226 230  
FT STRAND 231 239  
FT TURN 241 242  
FT STRAND 245 250  
FT TURN 251 252  
FT HELIX 254 257  
FT TURN 258 259  
FT STRAND 260 260  
FT STRAND 263 263  
FT HELIX 266 268  
SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031AA2DB CRC64;  
  
Query Match  
Best Local Similarity 31.3%; Score 305; DB 1; Length 290;  
Matches 65; Conservative 45; Mismatches 56; Indels 16; Gaps 4;  
  
QY 10 INFTAGATQVQSYTNFIRAVRGRLTLPNVRGLPI-----NORFILVELSNHAELSVT 62  
DB 25 VSPFLSGATTISYGVFFIKNLREALPYERKYNIPLLRSLSSGSGRYTLHLHTYADETIS 84,  
QY 63 LALDVTNAYVVGVRAGNSAVFFHFDNOEDA-EAITHLFTDVQNRNRYTFAFGNDRLEQLA 121

Db 85 VADVTVNYMGVAGDVSYFF---NEASATEAKVFVDKAKKVTLPYSGNYERLQTAA 141  
QY 122 GNLRNIELNGPLBEAISALYYIYSTGGTOLPTLARSFFIICIMISEAARFOYIEGMRT 181  
Db 142 GKIRENIPGLPALDSAITTLYYYTAS-----SAASALLVLIQSTAESARYKFIQQIGK 196  
QY 182 RI 183  
Db 197 RV 198  
  
RESULT 10  
ABRA ABRPR  
ID ABRA ABRPR STANDARD; PRT; 528 AA.  
AC P11140; P28589;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-a B chain].  
OS Abrus precatorius (Indian licorice) (Crab's eye).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
OX NCBI\_TaxID=3816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93132798; PubMed=8421313;  
RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;  
RT "Primary structure of three distinct isoabrinins determined by cDNA sequencing. Conservation and significance.";  
RL J. Mol. Biol. 229:263-267(1993).  
RN [2]  
RP SEQUENCE OF 1-251.  
RC TISSUE=Seed;  
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;  
RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from the seeds of Abrus precatorius.";  
RL Agric. Biol. Chem. 52:1095-1097(1988).  
RN [3]  
RP SEQUENCE OF 1-251 FROM N.A.  
RC TISSUE=Leaf;  
RX MEDLINE=91201329; PubMed=2016300;  
RA Evensen G., Mathiesen A., Sundan A.;  
RT "Direct molecular cloning and expression of two distinct abrin A-chains.";  
RL J. Biol. Chem. 266:6848-6852(1991).  
RN [4]  
RP SEQUENCE OF 262-528.  
RX MEDLINE=92371656; PubMed=1505674;  
RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;  
RT "The complete primary structure of abrin-a B chain.";  
RL FEBS Lett. 309:115-118(1992).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).  
RX MEDLINE=95333188; PubMed=7608980;  
RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;  
RT "Crystal structure of abrin-a at 2.14 A.";  
RL J. Mol. Biol. 250:354-367(1995).  
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. ABRIN-A IS MORE TOXIC THAN RICIN.  
CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-



CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M98344; AAA32624.1; ALT INIT.  
CC EMBL; X54872; -; NOT ANNOTATED\_CDS.  
CC PIR; S32429; TZLSA.  
CC .PDB; 1ABR; 07-FEB-95.  
CC InterPro; IPR000772; Ricin\_B\_lectin.  
CC InterPro; IPR001574; RIF.  
CC Pfam; PF00852; Ricin\_B\_lectin; 6.  
CC Pfam; PF00161; RIF; 1.  
CC PRINTS; PR00396; SHIGARICIN.  
CC SMART; SM00458; RICIN; 2.  
CC PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
CC PROSITE; PS00275; SHIGA\_RICIN; 1.  
CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.  
FT CHAIN 1 251 ABRIN-A A CHAIN.  
FT PEPTIDE 252 261 LINKER PEPTIDE.  
FT CHAIN 262 528 ABRIN-A B CHAIN.  
FT DOMAIN 273 400 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 403 527 RICIN B-TYPE LECTIN 2.  
FT REPEAT 283 325 1-ALPHA.  
FT REPEAT 326 366 1-BETA.  
FT REPEAT 369 401 1-GAMMA.  
FT REPEAT 414 449 2-ALPHA.  
FT REPEAT 453 492 2-BETA.  
FT REPEAT 495 528 2-GAMMA.  
FT ACT\_SITE 164 164 BY SIMILARITY.  
FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 286 305 BY SIMILARITY.  
FT DISULFID 329 346 BY SIMILARITY.  
FT DISULFID 417 430 BY SIMILARITY.  
FT DISULFID 456 473 BY SIMILARITY.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).  
FT CONFLICT 202 202 MISSING (IN REF. 2).  
FT CONFLICT 298 298 N -> Y (IN REF. 4).  
FT CONFLICT 427 427 M -> L (IN REF. 4).  
FT CONFLICT 467 467 T -> P (IN REF. 4).  
FT CONFLICT 483 483 V -> L (IN REF. 4).  
FT STRAND 5 8  
FT TURN 10 11  
FT HELIX 14 28  
FT STRAND 32 33  
FT TURN 34 35  
FT STRAND 36 38  
FT TURN 42 43  
FT HELIX 47 49  
FT STRAND 51 57  
FT STRAND 63 69  
FT TURN 70 72  
FT STRAND 75 79  
FT TURN 83 86  
FT STRAND 88 89  
FT TURN 92 93  
FT HELIX 94 97  
FT TURN 100 101  
FT STRAND 103 106  
FT TURN 113 114  
FT HELIX 115 119  
FT TURN 124 126  
FT STRAND 129 131  
FT HELIX 131 142

FT TURN 143 144  
FT HELIX 148 167  
FT STRAND 168 188  
FT TURN 181 182  
FT STRAND 185 185  
FT HELIX 189 196  
FT TURN 197 197  
FT HELIX 198 207  
FT STRAND 212 220  
FT TURN 222 223  
FT STRAND 226 231  
FT TURN 232 233  
FT HELIX 235 239  
FT STRAND 240 240  
FT STRAND 243 243  
FT STRAND 248 248  
FT STRAND 268 268  
FT STRAND 276 277  
FT STRAND 279 280  
FT HELIX 282 284  
FT STRAND 286 289  
FT HELIX 290 292  
FT TURN 296 297  
FT STRAND 299 303  
FT HELIX 311 313  
FT STRAND 315 317  
FT TURN 319 320  
FT STRAND 322 325  
FT TURN 326 327  
FT STRAND 328 332  
FT TURN 337 338  
FT STRAND 340 344  
FT TURN 346 348  
FT HELIX 351 353  
FT STRAND 355 355  
FT STRAND 357 358  
FT TURN 360 361  
FT STRAND 364 366  
FT TURN 367 370  
FT STRAND 371 374  
FT TURN 380 381  
FT STRAND 383 383  
FT STRAND 385 387  
FT HELIX 393 395  
FT STRAND 398 399  
FT STRAND 406 408  
FT STRAND 410 411  
FT HELIX 413 415  
FT STRAND 417 421  
FT TURN 422 423  
FT STRAND 424 428  
FT TURN 432 433  
FT HELIX 435 437  
FT STRAND 439 441  
FT TURN 443 444  
FT STRAND 447 449  
FT TURN 450 451  
FT STRAND 452 459  
FT TURN 464 465  
FT STRAND 467 472  
FT TURN 474 475  
FT HELIX 478 480  
FT STRAND 483 484  
FT TURN 486 487  
FT STRAND 490 492  
FT TURN 493 496  
FT STRAND 497 501

Query Match 31.1%; Score 303.5; DB 1; Length 528;  
Best Local Similarity 41.1%; Pred. No. 8.6e-21;  
Matches 76; Conservative 26; Mismatches 64; Indels 19; Gaps 5;



```

RESULT 12
RIPI_CUCFI
ID RIPI_CUCFI STANDARD; PRT; 286 AA.
AC C9FRX4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)
DE (EC 3.2.2.22).
DE Cucumis figareii.
OC Cucurbitaceae;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```

```

[1]
RN      SEQUENCE FROM N.A.
RP      Yamada T., Ohki S.T., Osaki T.;
RT      "Cloning and analysis of a cDNA coding a putative ribosome-
RT      inactivating protein from Cucumis figarei.";
RL      Plant biotechnol. 17:337-340(2000).
CC      -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC      specific adenosine on the 28S rRNA.
CC      -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC      TYPE 1 RIP SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC      or send an email to license@isb-sib.ch).

```

	Matches	60;	Conservative	46;	Mismatches	64;	Indels	15;	Gaps	3
QY	10	INFTTAGATGVSQTNFIRAVRGRLT <sup>1</sup> VLPN <sup>2</sup> RVLPI-----NQRFILVELSNHAELSVT	62							
Dd	28	VKFSLLGSNHKSYSKFITS <sup>3</sup> KRNALPNAGDIYN <sup>4</sup> TPLLVPSISGRRYLMQLMSNYEGNIIT	87							
QY	63	LALDVNTVNVGVYGAGNSAYFFHPD <sup>5</sup> NQEDAEA <sup>6</sup> THLFTVDVNQRNYTFAGCNYYDRLEQLAQ	122							
Dd	88	MAVDVTNNYIMGYLVNGTSYFF---NETDAQASKSEVFQCTKSITL <sup>7</sup> PYSGNYQKLQSVAR	144							
QY	123	NLRNIELGNGLPEEAISALYYSTGTOLPTLARSFIIICIQISEAARFOYIEGEMRTR	182							
Dd	145	KERSIPLGFNALDSALISTLYYYDSRSAPI-----AFLVLIQTTEAAARYKIETKOIIDR	199							
QY	183	IIRYNR	187							
Dd	200	ISVSK	204							

AC P16094; P24697;  
AD 01-APR-1990 (Rel. 14, Created)  
AE 01-MAR-1992 (Rel. 21, Last sequence update)  
AF 15-MAR-1992 (Rel. 42, Last annotation update)  
AG Ribosome-inactivating protein momordin I precursor (rRNA  
AH N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).  
AI Momordica charantia (Bitter melon) (Balsam pear).  
AJ Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AK Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
AL eucosids I; Cucurbitales; Cucurbitaceae; Momordica.  
AM NCBI\_TaxID=3673;  
AN [1]  
AO SEQUENCE FROM N.A.  
AP RIBOSOME-INACTIVATING PROTEIN  
AQ TISSUE=Seed;  
AR MEDLINE=91159486; PubMed=2001404;  
AS Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;  
AT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating  
AU protein.";  
AV Biochim. Biophys. Acta 1088:311-314(1991).  
AW [2]  
AX SEQUENCE OF 24-38.  
AY TISSUE=Seed;  
AZ MEDLINE=89326691; PubMed=2753596;  
BA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Sorio M.,  
BB Lappi D.;  
BC "N-terminal sequence of some ribosome-inactivating proteins.";  
BD Int. J. Pept. Protein Res. 33:263-267(1989).  
BE [3]  
BF SEQUENCE OF 24-70.  
BG TISSUE=Seed;  
BH MEDLINE=89005108; PubMed=3262509;  
BI Casellas P., Dussossoy D., Falasca A.I., Barbieri L., Stirpe F.,  
BJ Guillemot J.C., Ferrara P., Bolognesi A., Cennini P., Stirpe F.;  
BK "Trichokirin, a ribosome-inactivating protein from the seeds of  
BL Trichosanthes kirilowii Maximowicz. Purification, partial  
BM characterization and use for preparation of immunotoxins.";  
BN Eur. J. Biochem. 176:581-588(1988).  
BO [4]  
BP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
BQ MEDLINE=94356447; PubMed=8075985;  
BR Ren J., Wang Y., Dong Y., Stuart D.I.;  
BS "The N-glycosidase mechanism of ribosome-inactivating proteins  
BT implied by crystal structures of alpha-momorcharin.";  
BU Structure 2:7-16(1994).  
BV [5]  
BW X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).  
BX MEDLINE=94192822; PubMed=8143869;  
BY Husain J., Tickle I.J., Wood S.P.;  
BZ "Crystal structure of momordin, a type I ribosome inactivating  
CA protein from the seeds of Momordica charantia.";  
CB FEBS Lett. 342:154-158(1994).  
CC [6]  
CD X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
CE MEDLINE=95344383; PubMed=7619070;  
CF Huang Q., Liu S., Tang Y., Jin S., Wang Y.;  
CG "Studies on crystal structures, active-centre geometry and  
CH depurinating mechanism of two ribosome-inactivating proteins.";  
CI Biochem. J. 309:285-298(1995).  
CJ -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CK specific adenosine on the 28S rRNA.  
CL -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CM TYPE 1 RIP SUBFAMILY.  
CN -----  
CO This SWISS-PROT entry is copyright. It is produced through a collaboration  
CP between the Swiss Institute of Bioinformatics and the EMBL outstation  
CQ at the European Bioinformatics Institute. There are no restrictions on its  
CR use by non-profit institutions as long as its content is in no way  
CS modified and this statement is not removed. Usage by and for commercial  
CT entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CU or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CV -----  
CW EMBL; X57682; CAA40869.1; -;  
CX PIR; S14273; RLPUUG.  
CY -----  
CZ



Db 141 KIRENIDLGPLSSAITTLFYNA-----QSAPSAIIVLIQTTAAARFKYIERHVAKY 195

QY 183 IRYN 186

Db : |

196 VATN 199

Search completed: February 10, 2004, 16:23:26  
Job time : 6.33803 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 25.3521 Seconds  
(without alignments)  
1933.961 Million cell updates/sec

Title: US-10-083-336A-11

Perfect score: 975

Sequence: 1 MFEFQKYPINFTTAGATVQ.....RFQYIEGEMRTIRYNRRA 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*

- 1: sp archaea.\*
- 2: sp bacteria.\*
- 3: sp fungi.\*
- 4: sp human.\*
- 5: sp invertebrate.\*
- 6: sp mammal.\*
- 7: sp mhc.\*
- 8: sp organelle.\*
- 9: sp phage.\*
- 10: sp plant.\*
- 11: sp rodent.\*
- 12: sp virus.\*
- 13: sp vertebrate.\*
- 14: sp unclassified.\*
- 15: sp virus.\*
- 16: sp bacteriap.\*
- 17: sp archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	955	97.9	541	10 Q41174	Q41174 ricinus com
2	374.5	38.4	580	10 Q94BW3	Q94BW3 cinnamomum
3	372.5	38.2	580	10 Q94BW4	Q94BW4 cinnamomum
4	370.5	38.0	581	10 Q94BW5	Q94BW5 cinnamomum
5	368.5	37.8	549	10 Q9FV22	Q9FV22 cinnamomum
6	334	34.3	289	10 Q94KE4	Q94KE4 trichosan
7	332	34.1	289	10 Q41216	Q41216 trichosan
8	329.5	33.8	563	10 Q04367	Q04367 sambucus ni
9	326	33.4	247	10 Q9LRE3	Q9LRE3 trichosan
10	323.5	33.2	564	10 Q9AVR2	Q9AVR2 sambucus eb
11	318.5	32.7	528	10 Q06076	Q06076 abrus preca
12	317	32.5	270	10 Q8LPV7	Q8LPV7 trichosan
13	310.5	31.8	252	10 Q38760	Q38760 abrus preca
14	307.5	31.5	563	10 Q94S82	Q94S82 sambucus ni
15	307.5	31.5	563	10 Q8GT32	Q8GT32 sambucus ni
16	305	31.3	270	10 Q41611	Q41611 trichosan

17	303.5	31.1	252	10 Q38761	Q38761 abrus preca
18	302	31.0	565	10 Q04071	Q04071 sambucus ni
19	300.5	30.8	278	10 Q00980	Q00980 luffa cylin
20	298.5	30.6	547	10 Q9M6E9	Q9M6E9 abrus preca
21	297.5	30.5	251	10 Q96236	Q96236 abrus preca
22	296.5	30.4	251	10 Q96237	Q96237 abrus preca
23	293.5	30.1	566	10 Q04072	Q04072 sambucus ni
24	290.5	29.8	251	10 Q96235	Q96235 abrus preca
25	273	28.0	264	10 Q9FSH2	Q9FSH2 momordica c
26	273	28.0	570	10 Q41358	Q41358 sambucus ni
27	271	27.8	592	10 Q8W2E7	Q8W2E7 iris hollan
28	270	27.7	249	10 Q8LKQ5	Q8LKQ5 viscum albu
29	270	27.7	286	10 Q9FUW7	Q9FUW7 momordica c
30	267	27.4	570	10 Q22415	Q22415 sambucus ni
31	266.5	27.3	604	10 Q9M654	Q9M654 polygonatum
32	265	27.2	286	10 Q41257	Q41257 momordica c
33	263	27.0	565	10 Q8W243	Q8W243 viscum albu
34	260	26.7	254	10 Q8LKQ6	Q8LKQ6 viscum albu
35	260	26.7	573	10 Q8W2E8	Q8W2E8 iris hollan
36	259	26.6	251	10 Q8LKQ4	Q8LKQ4 viscum albu
37	259	26.6	293	10 Q8S452	Q8S452 jatropa cu
38	259	26.6	569	10 Q93543	Q93543 sambucus ni
39	253	25.9	531	10 Q8RXH6	Q8RXH6 viscum albu
40	251	25.7	249	10 Q8RXH7	Q8RXH7 viscum albu
41	250	25.6	258	10 Q9S9E4	Q9S9E4 gelonium mu
42	247	25.3	293	10 Q8VYU0	Q8VYU0 jatropa cu
43	246.5	25.3	203	10 Q8RY69	Q8RY69 gynostemma
44	246.5	25.3	275	10 Q8H1Y4	Q8H1Y4 gynostemma
45	245.5	25.2	603	10 Q9M653	Q9M653 polygonatum

## ALIGNMENTS

## RESULT 1

Q41174 ID Q41174 PRELIMINARY; PRT; 541 AA.  
AC Q41174;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)  
DE Proridin A chain (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
DE Ricinus communis (Castor bean).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92338377; PubMed=1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT "Molecular cloning of ricin."  
RL Targeted Diagn. Ther. 7:81-97(1992).  
CC -|- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC -|- SPECIFIC ACTIVITY: BELONGS TO THE 28S rRNA.  
CC -|- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; S40366; AAB22582.1; -.  
DR HSSP; P02879; 1BR6.  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR InterPro; IPR001400; Somatotropin.  
DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
KW Hydrolase; Toxin.  
FT NON TER 1  
SQ SEQUENCE 541 AA; 60281 MW; 2B7B2CDEF1F2E9D9 CRC64;

```

Query Match          97.9%; Score 955; DB 10; Length 541;
Best Local Similarity 95.0%; Pred. No. 1.3e-82;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFARG 111
DB 61 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFARG 120

QY 112 GNYDRLEQLAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAAR 171
DB 121 GNYDRLEQLAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAAR 180

QY 172 FQYIEGEMTRIRNRSR 190
DB 181 FQYIEGEMTRIRNRSR 199

RESULT 2
Q94BW3 PRELIMINARY; PRT; 580 AA.
AC Q94BW3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin III precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
DR EMBL; AY039803; AAK82460.1; -
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS0231; RICIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32
FT CHAIN 33 580
FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN III.
SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match          38.4%; Score 374.5; DB 10; Length 580;
Best Local Similarity 48.1%; Pred. No. 3.9e-27;
Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTTKNATKTSYTFEALRAQLASGEHPHGVPRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAGGNYDR 116
DB 93 ADSPVTLAVDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAGGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGERREELIGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEVRFRFIE 207

Query Match          38.2%; Score 372.5; DB 10; Length 580;
Best Local Similarity 47.6%; Pred. No. 6.1e-27;
Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTTKNATKTSYTFEALRAQLASGEHPHGVPRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAGGNYDR 116
DB 93 ADSPVTLAVDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAGGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGERREELIGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEVRFRFIE 207

```

```

QY 177 GEMRTRI 183
DB 208 YRVRESI 214

RESULT 3
Q94BW4 PRELIMINARY; PRT; 580 AA.
AC Q94BW4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin II precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
DR EMBL; AY039802; AAK82459.1; -
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS0231; RICIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32
FT CHAIN 33 580
FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN II.
SQ SEQUENCE 580 AA; 64265 MW; 37E4289ECCE0CBFF CRC64;

Query Match          38.2%; Score 372.5; DB 10; Length 580;
Best Local Similarity 47.6%; Pred. No. 6.1e-27;
Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56
DB 33 YQVTVTTKNATKTSYTFEALRAQLASGEHPHGVPRERSTVPDSKRFILVLSNWA 92

QY 57 AELSVTALDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAGGNYDR 116
DB 93 ADSPVTLAVDVTNAYVGYRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAGGNYDR 149

QY 117 LEQAGNLRENELGNGLPGLAEISALYYSTGGTQLPTLARSFFICIQMISEAARFOYTE 176
DB 150 LERVAGERREELIGMDPLENAISALWISNL--NQORALARSLIVVIQWVAEVRFRFIE 207

QY 177 GEMRTRI 183
DB 208 YRVRESI 214

```

```

RESULT 4
Q94BW5 PRELIMINARY; PRT; 581 AA.
AC Q94BW5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin I precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).

```







```
CC -I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -I- SIMILARITY BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; AB039324; BAA92530.1; -.
DR HSP; P09989; IWRJ.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR DR PROSITE; PS00275; SHIGA_RICIN; 1.
DR KW Hydrolyase; Toxin.
FT NON_TER      1
FT TER          247
SQ SEQUENCE     247 AA; 27199 MW;  89811AC32892F03F CRC64;

Query Match              33.4%; Score 326; DB 10; Length 247;
Best Local Similarity    39.6%; Pred.No.5.2e-23;
Matches       72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

QY      10 INFETAGATVQSNTFNRAVR-----GRLTVLPLNRVGLPINORFILVELSNHAELSVT 62
DB      2 VSFELSGLATSSSGVFVISNLRKALPYERKLYDIPLLRLSTLPGSORYALIHLITVADETTIS 61
QY      63 LALDVTNAYVGVGRAGNSAYFFHFDNQEDA-EAITHLFDTVDQNRYTFAFGNGYDRLEQLA 121
DB      62 VALDVTVNVVMYGVRAGDTSYFP---NEASATEAAKYVKDAKRKVTLTPSYGNVERLQIAA 118
QY      122 GNLENILELGNGPLEEAISALESXYSTGGTQTLPILARSFCIIOMISAAARFOYIEGEEMRT 181
DB      119 GKRENIPLGLPALDSAITTLFYNNAN-----SAASLMVLIOGSTSAARYKFIEQQIGK 173
QY      182 RI 183
DB           |
DB      174 RV 175

RESULT 10
Q9AVR2 PRELIMINARY; PRT; 564 AA.
ID Q9AVR2 AC Q9AVR2;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Ribosome-inactivating protein precursor (EC 3.2.22) (rRNA
DE N-glycosidase).
DN EBUL.
GN Sambucus ebulus.
OC Okuyayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.
OX NCBI_TaxID=28503;
RN [1]
RP SEQUENCE FROM N.A.
RR TISSUE=Leaf;
RC Girbes T., Iglesias R., Perez Y., Ferreras J.M., Citores L.;
RA "Molecular cloning of ebulin 1.",
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -I- SIMILARITY BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; A400822; CAC33178.1; -.
DR HSP; P02879; 2AAI.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Dr Prosite; IPR001574; RIP.
DR Pfam; PF00652; Ricin B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Glycosidase; Hydrolase; Signal; Toxin.
FT SIGNAL      1 25 POTENTIAL.
FT CHAIN       26 298 EBULIN_L_A-CHAIN.
FT CHAIN       299 564 EBULIN_L_B-CHAIN.
```

RESULT 14	
Q945S2	PRELIMINARY; PRT; 563 AA.
ID	Q945S2
AC	Q945S2;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Ribosome-inactivating protein (EC 3.2.2.22) (IRNA
DE	N-glycosidase).
GN	AVL.
OS	Sambucus nigra (European elder).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC	Asteridae; Camanulids; Dipsacales; Adoxaceae. Sambucus

OX	NCBI_TaxID=4202;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Leaf;
RA	Van Damme E.J.M.;
RT	"Characterization and cloning of lectins and ribosome-inactivating
RL	proteins from Sambucus nigra leaves.";
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC	-I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.
CC	-I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR	EMBL: AF409135; AAL04123.1; -
DR	InterPro: IPR000772; Ricin_B_lectin.
DR	InterPro: IPR001574; RIP.
DR	Pfam: PF00652; Ricin_B_lectin; 6.
DR	Pfam: PF00161; RIP; 1.
DR	PRINTS: PR00396; SHIGARICIN.
DR	SMART: SM00458; RICIN; 2.
DR	PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR	PROSITE: PS00275; SHIGA_RICIN; 1.
KW	Hydrolase; Toxin.
SQ	SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;

Query Match	31.5%;	Score 307.5;	DB 10;	Length 563;
Best Local Similarity	38.4%;	Pred. No. 9.5e-21;		
Matches	73;	Conservative	36;	Mismatches 56;
				Indels 25;
				Gaps 5;

Qy	7	YPINFTTAGATVQSYNFI	TRAVR-----GRLTVPENRVGLPFI	QINQFI	VELSNH	56	
Db	28	YPSVSFLNDGAKSATYR	DFLSNRKTVATGT	YEVNGLPVLRRSE	SEVQKSFVLVPLTNY	87	
Qy	57	AELSVTALDVTNAYVY	GYPAGNSAYFFHP	PNQDAEAI--THLFTDVQNYRT	YAFGQNY	114	
Db	88	NGMTVTLAVDVTNLV	YVAFSGNANSYFF-----KDA	TEVQKSNLFPVGTQKQ--TL	SFTGNY	141	
Qy	115	DRLEQLAGNLRN	ELGNGPLEEAL	SALYYYSTGGTQ	PTLARBFIIC	OMISEARFQY	174
Db	142	DNLETAANTRES	ELGSPDLG	AITSLYHGD-----SVARS	LLVVIQWVS	AAARFY	194
Qy	175	IEGEMRTRIR	184				
Db	195	IEOEVRRESLO	204				

```

RESULT 15
Q8GT32
ID Q8GT32 PRELIMINARY; PRT; 563 AA.
AC Q8GT32;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE type 2 ribosome-inactivating protein nigrin 1 precursor
DE (EC 3.2.2.22).
DE Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asterales; Campanulids; Dipsacales; Adoxaceae; Sambucus.
OC NCBI_TaxID=4202;
RN [1]
SEQUENCE FROM N.A.
RP
RC TISSUE=Leaf;
RA
RA Gibbs T., Arias F.J., Antolín P.;
RT "Characterization and molecular cloning of Nigrin 1, a type two
RT ribosome-inactivating protein from leaves of elder (Sambucus nigra).";
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF249280; AA086130.1; -
DR Hydroxylase; Glycosidase.
SQ
SQ SEQUENCE 563 AA; 621173 MW; 0EB236421FC5E04F CRC64;

```

Query Match	31.5%;	Score 307.5;	DB 10;	Length 563;
Best Local Similarity	38.4%;	Pred. No. 9.5e-21;		
Matches 73;	Conservative	36;	Mismatches 56;	Indels 25;
Gaps	5;			

```

QY      7 YPIINFTTAGATVQSYTNFIBAVR-----GRUTVLPNRVGLPFINORFILVELSNH 56
Db      28 YPSVFNLDGAKSATYRDFLSNLKRTVATGYEVNGLPVLRRSEVQVKSRFVLVPLTNY 87
QY      57 AELSVTALADVTNAAVVGYRAGNSAYFFHPDNORDAEAI--THLFTDVQNRYVFAFGNY 114
Db      88 NGNTVTILAVDVTNLYVAFSGNANSYFF-----KDTEVOKSNLFTVGTGKN-TLSTGTNY 141
QY      115 DLKLOLAGNLRENTLGNGLPBLEEIASLYYSTGGTOLPTLARSFFICIMISEAARFOY 174
Db      142 DNLETAANTRRSESTELGSPDLGTAITSLYHGD-----SVARSLLWQMVSEAAARFY 194
QY      175 IEGEMRTRIR 184
Db      195 IEQEVRSLQ 204

```

Search completed: February 10, 2004, 16:26:40  
Job time : 26.3521 secs

Search completed: February 10, 2004, 16:26:40  
Job time : 26.3521 secs